Stimulus Representation and Processing in Human Associative Learning:

An evaluation of current elemental and configural associative learning theories

Dissertation

zur

Erlangung des Doktorgrades

der Naturwissenschaften

(Dr. rer. nat.)

dem

Fachbereich Psychologie

der Philipps-Universität Marburg

vorgelegt von

Anna Barbara Thorwart

aus Nairobi, Kenia

Marburg, Lahn 2009

Vom Fachbereich Psychologie der Philipps-Universität Marburg als Dissertation am _____ angenommen.

Erstgutachter: Prof. Dr. Harald Lachnit, Philipps-Universität Marburg

Zweitgutachter: Prof. Dr. John Pearce, Cardiff University

Tag der mündlichen Prüfung am: 22. Januar 2010

Acknowledgement

I would like to thank all those who supported me during preparation of this thesis.

First and foremost I offer my sincerest gratitude to my supervisor, Prof. Harald Lachnit, who has supported and guided me throughout my thesis with knowledge and humour whilst allowing me the room to work in my own way. I would also like to thank Prof. John Pearce, who was so kind to agree to be my external supervisor and international mentor before even knowing myself or my project.

A special thanks to Dr. Steven Glautier for the time I had the pleasure to spend with him in Southampton and for his lasting support.

This thesis would not have been possible without my colleagues: Dr. Anja Lotz, who endured three-years of questions and comments, Dr. Metin Üngör, who always had a sympathetic ear and proper solutions, and Stephan König, for introducing me to Matlab and a lot of other stuff. And of course a special thanks to all our student assistants that collected the data.

Last but not least, I am grateful to the Deutsche Forschungsgemeinschaft (DFG) graduate program “NeuroAct” for their support.

Zusammenfassung

Assoziative Lerntheorien beschreiben das Erlernen von Zusammenhängen zwischen Ereignissen, zum Beispiel zwischen dem Essen eines Apfels und darauf folgenden Bauchschmerzen. Lange Zeit wurden dabei zwei Klassen von Modellen gegenübergestellt. Während in elementaren Theorien Reizkonfigurationen in ihre Komponenten zerlegt werden und diese dann jeweils Assoziationen eingehen, werden in konfiguralen Modellen Reizkonfigurationen als Ganzes miteinander verknüpft. Aufgrund der widersprüchlichen empirischen Befundlage stellt sich jedoch die Frage, ob Menschen und Tiere nicht in der Lage sind beide Arten der Reizrepräsentation anzuwenden. Demgegenüber postulieren moderne Theorien wie die elementaren Modelle von Wagner (2003) und Harris (2006) oder das erweiterte konfigurale Modell von Pearce (Kinder & Lachnit, 2003) Flexibilität innerhalb elementarer bzw. konfiguraler Grundprinzipien und stellen die Existenz einer zweiten grundsätzlich unterschiedlichen Form der Reizrepräsentation in Frage. Um diese beiden Ansätze zu vergleichen, wurden Vorhersagen der elementaren und konfiguralen Modelle gegenübergestellt. Hierzu wurde zunächst ein benutzerfreundliches Computerprogramm zur Simulation der Modelle geschrieben. Daraufhin wurden in zwei Serien von Lernexperimenten mit Menschen sowohl sich widersprechenden Vorhersagen empirisch getestet als auch Faktoren des experimentellen Settings manipuliert, um so die Art der Reizverarbeitung zu beeinflussen und ihre Wirkmechanismen zu untersuchen.

Insgesamt machen die Ergebnisse aller Experimente deutlich, dass auch moderne elementare Modelle konfigurale Modelle nicht ersetzen können. Stattdessen stimmte das beobachtete Verhalten sowohl aus der Lernphase zweier Feature-Negative Diskriminationen als auch die Generalisationsdekremente nach erfolgreichem Diskriminationstraining nur mit der erweiterten Form des konfiguralen Modells von Pearce überein. Im Kontext bisheriger Befunde spricht dies für die Existenz zweier Arten der Reizrepräsentation. Jedoch konnten keine eindeutigen Einflussfaktoren nachgewiesen werden, die bestimmen, welche Art der Repräsentation Menschen in einer Situation wählen. Weder Manipulationen des Reizmaterials (perzeptuelle Gruppierung durch Bewegung, räumliche Anordnung der Reize) noch der experimentellen Durchführung (Lernparadigma, Zeitdruck während der Reizdarbietung, kausale Charakter des Szenarios) zeigten Effekte. Stattdessen wurde deutlich, dass zukünftige Forschung sich auf die Umstände, die die Bedeutung der untersuchten Faktoren kontrollieren, konzentrieren muss. Abschließende theoretische Analysen verdeutlichen darüber hinaus, dass die beobachtete Überlegenheit des konfiguralen Modells nicht auf der konfiguralen Repräsentation an sich beruht, sondern auf eine damit einhergehende Normalisierung der Aktivierungsstärke zurückgeht. Daraus folgernd sollten elementare Modelle, die diesen Mechanismus beinhalten, diesen Vorteil des konfiguralen Modells ausgleichen können.

Summary

Theories of associative learning describe learning about the relationship between two events, e.g. the eating of an apple and subsequent stomach ache. One important classification of these models is based on the stimulus representation they suppose. Whereas elemental models assume that the representations of a stimulus compound consist of representations of its components establishing associations, configural models propose that stimulus compounds are represented and associated as a whole. However, as the empirical results have not consistently favoured one class of models, it was suggested that humans and animals can switch between both modes of stimulus representation. Alternatively, modern theories as the elemental model of Wagner (2003) and Harris (2006) or the extended configural model of Pearce (Kinder & Lachnit, 2003) postulate flexibility within either elemental or configural principles and question the existence of a second, fundamentally different mode of stimulus representation. In order to evaluate these two hypotheses, we contrasted the predictions of the elemental and configural models. Firstly, a user-friendly environment for computer simulation of the models was written. Then, contradicting predictions were tested in two series of human learning experiments. Furthermore, it was investigated whether manipulations of the experimental setting would influence stimulus processing and on which mechanism these effects are based.

The results demonstrate that models that comprise configural principles are not made superfluous by modern elemental models. Instead, the extended version of Pearce's configural model was able to account for all behaviour observed both during learning of two feature-negative discriminations as well as in generalisation tests after learning. Together with the "elemental" results of other studies, this corroborates the position that there are two modes of stimulus processing during associative learning. Further tests and experiments, however, are necessary concerning factors that influence which mode is chosen in a learning situation. Manipulations of the stimulus material (perceptual grouping by motion; spatial arrangement) as well as manipulations of the experimental procedure (learning paradigm; time pressure during stimulus presentation; causal nature of the cover story) had no effects. This indicates that more controlled research is needed investigating not only the factors but also the circumstances controlling these factors' relevance. Furthermore, additional theoretical analyses revealed that the observed advantage of the configural model is not based on the configural representation itself but on a normalisation of the representation's activation. Therefore, elemental models incorporating this mechanism could compensate their present shortcoming.

Content

Proposed Manuscripts

I. Thorwart, A., Schultheis, H., König, S. & Lachnit, H. (2009). ALTSim: A MATLAB simulator for current associative learning theories. *Behavior Research Methods*, 41, 29-34.

II. Thorwart, A., Glautier, S., & Lachnit, H. (2009). *Convergent results in eyeblink conditioning and contingency learning in humans: Addition of a common cue does not affect feature-negative discriminations.* In preparation for submission to Biological Psychology.

III. Thorwart, A. & Lachnit, H. (2009). Symmetrical generalisation decrements: Configural stimulus processing in human contingency learning. *Learning & Behavior*, 37, 107-115.

IV. Thorwart, A. & Lachnit, H. (2009). *Generalization decrements and salience: new data and more elaborated analyses.* Submitted for publication to Learning & Behavior.

Introduction

For humans, as for all other animals, learning about the relationship between events and stimuli in their environment is crucial for survival. It is necessary in order to predict future events, harmful and enjoyable, based on the perception and processing of present events. Hence, the knowledge gained through such learning is the foundation for adapting our behaviour to the environment and preparing for the future, altering it or even impeding it.

Within psychological research, the classical domains for observing such learning are conditioning, predictive and causal learning (Dickenson, 2001), but also spatial learning (Pearce, 2009) or categorisation (Gluck & Bower, 1988) can be understood as learning about relationships. Furthermore, its concepts are applied for example to the development of prejudices (Walther, Nagengast, & Trasselli, 2005) and are prominent in models of anxiety and panic disorders (Bouton, Mineka, & Barlow, 2001). Despite of this prevalence, the research of the last century could not figure out what exactly happens when human and nonhuman animals learn about the relationship between two events.

A main theoretical approach conceptualises such learning as establishing and changing associative links between mental representations of the predictive cues (or conditioned stimuli, CS, in terms of conditioning) and the outcome (i.e. unconditioned stimulus, US). The strength (or weight) of an association then reflects the learned relationship between the events. However, different models of associative learning focus on different aspects (see Pearce & Bouton, 2001): whereas some specify what happens with the associability of the stimuli during training (Pearce & Hall, 1980; Mackintosh, 1975), other models, e.g. Wagner's SOP model (Wagner, 1981), are focussing on within-trial processes; Miller's comparator hypothesis, in turn, concentrates on the behavioural expression of the associations (Miller & Matzel, 1988). The present thesis is concerned, however, with models describing in detail the structure of the internal representation of the predictive cues.

An appropriate description of their representation is central to understand fundamental learning processes like stimulus generalisation and discrimination. Stimulus generalisation describes the phenomenon that a learned response is not only shown in answer to the previously trained stimulus but also when we are confronted with a new, similar stimulus. Discrimination describes the somewhat opposite process by which we learn to respond differentially to several stimuli. For example, a behaviour therapist has to ensure that his patient generalises a learned behaviour from the therapeutic setting to everyday situations and at the same time that the patient discriminates between situations where this behaviour is appropriate or inappropriate. Models of associative learning explain generalised responding to an apparently new stimulus by assuming that this new stimulus shares features with the previously trained stimulus. Discrimination learning instead relies on features in which the stimuli vary. As both processes are based on features and properties of stimuli, the knowledge of how these stimuli and their features are represented within the learning system is decisive in order to predict and explain stimulus generalisation and discrimination.

However, a part/whole or elemental/configural debate exists concerning stimulus representation within associative learning comparable to that in other domains of psychological research. In the following, an introduction to two basic models is given, the elemental theory of Rescorla and Wagner (1972) and the configural model of Pearce (1994). Then, the three models being investigated in the present thesis are described: the Replaced Elements Model of Wagner (2003, 2008), the elemental model of Harris (2006) and an extension of the configural Pearce Model by Kinder and Lachnit (2003).

1. The Nature of Stimulus Representation

The representations of the predictive stimuli determine the possible associations within an associative network as they constitute one side of the associative link and thus one argument of the rules that govern the associations' interactions during learning and performance. The following models all assume that learning follows the delta-rule as described for example in the theory of Rescorla and Wagner (1972).

Learning, i.e. changes to the associative strength, will happen as long as the prediction derived from the current associations of the predictive stimuli with the outcome does not equal the actual outcome. Associations established between predictive stimuli are not taken into account in the following. Instead, the focus is on the structure of representations that correspond to stimulus compounds consisting of several predictive stimuli and on their associations with the outcome.

Elemental models suppose that the representation of stimulus compounds comprises the representation of their components. In a simple linear elemental model like the Rescorla-Wagner Theory (Rescorla & Wagner, 1972), a stimulus compound AB excites the representations of its components, A and B, and each component will be associated individually with the outcome's representation. As the activation excited by each stimulus component then adds up to the activation excited by the stimulus compound, the associative strength of a compound is computed by the linear sum of its components' associative strengths and there is no additional information gathered about the compound itself. This was and is very successful in explaining learning effects like summation, blocking and other phenomena of cue competition, but not for nonlinear discrimination tasks like Negative Patterning (for a review see Miller, Barnet & Grahame, 1995). In nonlinear discriminations, the outcome of the stimulus compound does not equal the linearly summed outcome of the components themselves. In a Negative Patterning task for example, trials where the components are presented alone are followed by the outcome (A+, B+) but trials where the whole stimulus compound is presented are not (AB-). The Rescorla-Wagner Theory incorrectly predicts that such a discrimination cannot be learned because learning about the components in the reinforced trials generalises completely to the compound in the non-reinforced trials and vice versa. To overcome this limitation, an additional representational element that reflects special properties of the compound itself, a so called unique cue, was proposed (e.g., Whitlow & Wagner, 1972). The associative link of this unique cue is only activated when the compound is presented and could then mirror knowledge about the compound that differs from the summed knowledge that is gathered about the components themselves, i.e. the unique cue could counteract the

components' associations and their prediction of an outcome in the Negative Patterning task.

In contrast, a configural solution to this problem was proposed by Pearce (1987, 1994, 2002). His model supposes that the entire stimulus compound AB accrues to one configural unit "AB" and this unit enters into a single association with the outcome. A main problem of such a purely configural representation is that one would know nothing about the predictive value of a new compound, even if it consists for the most part of well known stimuli. For example, after learning something about A, this knowledge would not generalise to the stimulus compound AB because AB excites its own configural unit "AB" and not "A" and "B". This is functional for Negative Patterning tasks where the knowledge about the components would only interfere with learning about the compound, but not for most other tasks and real life situations. (Without generalisation, we would for example eat a fruit salad consisting of apples and bananas even if somebody told us that the apples are foul and, most likely, foulness of fruits does not follow a Negative Patterning schedule.) To circumvent this, configural units of other, currently not presented stimulus configurations (single stimuli and/or compounds) are assumed to be activated, too, but depending on their similarity to the actually presented stimulus configuration. The associations of the co-active configural units contribute in this way to the current prediction, i.e. the knowledge one has about their predictive power generalises to some degree. In most associative learning theories, similarity between two stimulus configurations is determined by the proportion of components common to both (Young & Wassermann, 2002). Following Pearce (1994), this can be captured as the square of the dot product of the normalised input vectors coding the presence or absence of the components or, equivalently, the square cosines of the angle these vectors draw in a multidimensional stimulus space.

$$S_{X,Y}=(\mathbf{o}_X\mathbf{o}_Y)^2=\cos^2(\mathbf{o}_X\mathbf{o}_Y) \quad (1)$$

$S_{X,Y}$ similarity between stimulus configuration X and Y

$\mathbf{o}_X$ normalised vector of input activation of configuration X

In order to maintain the configural character of the model, there is an additional winner-takes-all rule concerning learning so that it is only the association of the presented configuration's configural unit that is altered. Overall, the configural model is successful in handling nonlinear discriminations; however, because of the reduced generalisation between the configural units it has especially problems to correctly predict learning processes that are based on linear summation effects.

Even though these models make quite different predictions for particular experimental designs, the empirical results have not consistently favoured one class of models or the other. Instead, it has become clear that results in line with both ways of stimulus representation can be obtained, and several factors have been identified that influence the outcome of learning experiments (Melchers, Shanks, & Lachnit, 2008). As a result, flexibility concerning the nature of stimulus representation was requested as a move beyond an either-or approach (e.g., Glautier, 2008). One way to achieve this is to postulate two modes of stimulus representation, an elemental and a configural, and to assume that humans and animals can switch between them depending on situational factors. Another solution was proposed by the three models introduced in the following. They try to account for the contradicting results by introducing flexibility in representation following either elemental or configural principles.

To begin with modern elemental models, they pursue two goals: first they have to overcome the limitations of the simple linear elemental model of Rescorla and Wagner (1972), and second, they have to be able to account for the ambiguous results mentioned above. To account for the limitations for example concerning nonlinear discriminations, the Replaced Elements Model (Wagner, 2003, 2008) and the elemental model of Harris (2006) assume a nonlinear combination of stimulus components whenever they are presented together. More precisely, the representation of one component differs qualitatively in the Replaced Elements Model and quantitatively in the elemental model of Harris (2006) depending on whether the component is presented alone or in compound with other components (Livesey & Harris, 2008). Furthermore, both models comprise parameters controlling different properties of the representation and thus providing the necessary variability.

In the Replaced Elements Model (Wagner, 2003, 2008), the representation of a stimulus consists of several context-dependent and context-independent elements, each having its own associative link with the outcome representation. The context-independent elements of stimulus A (a_i) are always activated if A is present. By contrast, the activation of the context-dependent elements is not only determined by the presence of A itself but also by the presence of another stimulus, e.g. B. Some elements (a_b) are only activated if A is presented together with B, just as the unique cues proposed by Whitlow and Wagner (1972); other elements ($a_{\sim b}$), however, are only activated if A is present but not B. As the number of $a_{\sim b}$ and a_b elements is assumed to be equal, the latter replace the former whenever A is presented together with B so that the overall number of activated elements by A stays constant. It follows that similarity and generalisation between stimulus configurations depend not only on the number of common components and therefore commonly activated associations but also on the proportion of the context-dependent elements in each component's representation. With a proportion of 0 %, the representations of stimulus A and B consist only of the context-independent elements a_i and b_i and, keeping with the Rescorla-Wagner Theory, nothing will be replaced as these are the elements that are also excited by AB. Subsequently, the response to a stimulus compound AB is the same than to a previously trained A. In the case of a proportion of 100 %, all elements are context-dependent: there are no a_i and b_i that are activated both by AB and by A or B, respectively, and nothing is generalised. In the case of Negative Patterning task learning then occurs on completely different associations in the A+, B+ and AB- trials and solving is straightforward. Thus, the Replaced Elements Model overcomes the limitation of the Rescorla-Wagner Theory. To account for the second problem of the variable empirical results, the proportion of replaced elements is conceptualized as a free parameter, called *r*. A major determinant of the replacement parameter, *r*, is the different stimuli employed and the amount of perceptual interaction between two components. For many discrimination tasks, this parameter makes the model very powerful, as it is able to generate variable predictions, especially predictions that were previously only made by configural models (e.g., missing summation).

The elemental model of Harris (2006) supposes quantitative nonlinearity. As in the Replaced Elements Model, each component is represented by several elements that are associated individually. A compound is represented by the elements representing the individual components, and no additional information about the compound is processed. However, the amount of some elements' activation by the compound AB differs from their activation by A alone. At first, the activation depends on the salience of the elements. Highly salient elements (e.g., elements representing more intensive features) are highly activated. Then, elements compete against each other for entry in an attention buffer on the basis of the increase in their activation, and any element that captures attention receives a further boost to its activation. Because of this competition between the elements, some of the elements that are boosted when their components are presented alone are only weakly activated when their components are presented in compound. As the activation strength ultimately serves as learning parameter in the learning rule, there will be less learning about these elements in compound trials. Furthermore, if two components have some features in common, they activate the same elements twice. This increases the nonlinearity of the activation even further and enhances the model's capacity to solve nonlinear problems. The Harris Model is able to make some predictions that are not made by other elemental models, including the Replaced Elements Model, but by the configural model of Pearce (e.g., the influence of a common element on feature-negative tasks, see below). Yet, it is a purely elemental model. Furthermore, the possibility to control the different theoretical components like the capacity or the boost of the attention buffer introduces the necessary flexibility.

On the configural side, Kinder and Lachnit (2003) developed an extension of the Pearce Model. They did not assume a change to the kind and structure of the representation but introduced a discrimination parameter, d, that reflects the overall discriminability between stimulus configurations. This parameter replaces the exponent within the above introduced similarity equation of Pearce (1994; see Equation 1).

$$S_{X,Y}=(\mathbf{o}_X\mathbf{o}_Y)^d=\cos^d(\mathbf{o}_X\mathbf{o}_Y) \quad (2)$$

$S_{X,Y}$ similarity between stimulus configuration X and Y

$\mathbf{o}_X$ normalised vector of input activation of configuration X

d discrimination parameter

The final similarity computation used in Pearce (1994, see Equation 1) corresponds subsequently to d = 2. With higher values of d, the generalisation gradient becomes steeper, i.e. there is less generalisation independent of the proportion of common components. As generalisation is the central process in the Pearce Model, this enables the extended model to make some unique predictions (e.g., Lachnit, Schultheis, König, Üngör, & Melchers, 2008; Pearce, Esber, George, & Haselgrove, 2008).

Obviously, the flexibility of all these models makes them more powerful. But is one of them flexible enough to make others models superfluous? This would indicate that the flexibility observed in the results is not due to fundamentally different modes of stimulus processing and representations during learning but accountable within the framework of either elemental or configural processing. The first intention of the present thesis is to reassess if one of the models correctly predicts behaviour during learning as well as in tests after learning. Therefore, a discrimination and a generalisation task were chosen so that the models, despite their flexibility, make differing predictions that can be tested against each other.

The second aim is to investigate factors that influence stimulus processing and representation, either within one model or between an elemental and a configural mode. As mentioned above, Melchers et al. (2008) already identified some factors in their review. Most factors concern the circumstances of learning such as the prior experience (Melchers, Lachnit, & Shanks, 2004) or the cover story (Waldmann, 2007). Furthermore, the kind of stimulus material itself can also have an influence. In this line, Wagner (2003) proposed that the *r* parameter of his Replaced Elements Model may be controlled by perceptual interaction between the components and that perceptual

interaction is influenced for example by the modality of the components or whether they combine in an integral or separable way (see also Myers, Vogel, Shin & Wagner, 2001; Lachnit, 1988). Several studies also showed that different spatial arrangement of the stimulus components on the screen results in differing stimulus processing (Glautier, 2002; Livesey & Boakes, 2004). The present thesis further investigates these effects as well as their mechanisms.

Outline of the present thesis

The thesis is composed of four papers that will be described shortly in the following. The first paper presents an introduction to a simulation program for current associative learning theories (chapter 1). The remaining three papers cover experimental studies that were designed to evaluate the models' predictive power. The first of them describes experiments conducted to compare acquisition of two feature-negative problems (chapter 2.1) whereas the last two papers constitute the main part of the thesis and report on generalisation decrements after successful learning (chapter 2.2.1 and 2.2.2).

1. ALTSim: A MATLAB simulator for current associative learning theories.

To compare theories empirically, one has to find procedures and tasks for which their predictions of the observed behaviour differ. The central advantage of most associative theories since Rescorla and Wagner (1972), and probably one reason for their popularity, is that they can be specified as mathematical models that again can be implemented in a computer algorithm. Therefore, computer simulations are a central method in investigating associative learning (see for example Haselgrove, Robinson, Nelson, & Pearce, 2008). In order to enhance the ease of use and thereby facilitate the application, we programmed a MATLAB® based graphical user interface for simulating several current associative models.

The main principles and features are:

1. Units and attributes that are applied to all models have to be specified only once.

2. Simulations should be as realistic as possible. It is for example possible to randomize trial sequences or to model the empirically used trial sequences for multiple subjects.
3. Parameters are easy to access and to change.
4. The output should be comprehensive. Associative weights and resulting outcome activations can be produced and saved graphically as well as numerically for any stimulus at any time of learning.

2. Empirical studies

2.1. Convergent results in eyeblink conditioning and contingency learning in humans: Addition of a common cue does not affect feature-negative discriminations.

In order to evaluate the Rescorla-Wagner Theory, Pearce and Redhead (1993) investigated the influence of a common cue on a feature-negative discrimination. In a feature-negative discrimination, a stimulus A is followed by an outcome when presented alone but not when presented together with another stimulus B, the so-called feature-negative. In an autoshaping experiment, Pearce and Redhead compared this A+/AB- discrimination with an AX+/ABX- problem where a cue is added to both kinds of stimulus configurations. The first aim of the present study was to investigate acquisition of these two discrimination problems in two different paradigms of human associative learning, eyeblink conditioning as well as contingency learning. Despite previous replications of experiments of one paradigm in the other, this is the first time that both were tested in parallel, adapting the procedures as far as possible and reasonable. Secondly, the predictions of the Replaced Elements Model (Wagner, 2003) as well as the predictions of the extension of the Pearce Model (Kinder & Lachnit, 2003) were compared for these discriminations.

Due to the linear summation of activation and associative strength for compounds, the Rescorla-Wagner Theory predicts that the more stimuli are around the more will be learned. This is even independent of the relevance of the components for the discrimination learning and leads to the counterintuitive prediction that adding a common cue to two configurations will facilitate their discrimination, even if at the

same time adding a common cue to two configurations will increase their similarity. As mentioned above, the assumption of linear summation of activation is retained in the Replaced Elements Model. Hence, the Replaced Elements Model like the Rescorla-Wagner Theory predicts that response differentiation in feature-negative discriminations will be facilitated by adding a common cue to the reinforced and non-reinforced stimulus. In the Pearce Model, the activation of the configural units is weighted by their similarity to the presented configuration. The more similar two stimulus configurations are, the more similar will be the activation of configural units by them and the more their associative strength will generalise between them. This will impede discrimination learning. However, increasing the discriminability in the extension of Kinder and Lachnit (2003) will reduce generalisation in general and the influence of higher similarity will diminish.

In two experiments, participants received both feature-negative discriminations as within-subject manipulation. The first experiment was a human eyeblink conditioning study; in the second, a contingency learning paradigm was introduced where the air pollution produced by an airplane had to be predicted. Both experiments used the same coloured circles as predictive cues and both experiments obtained convergent results: Adding a common cue did not affect response differentiation at all. Furthermore, inhibitory learning of the feature-negative did not differ between discriminations, as tested additionally in contingency learning.

These results clearly contradict the predictions of the Replaced Elements Model and show its limitations. Instead, the extended Pearce Model with a high discrimination parameter ($d > 20$) successfully predicts the pattern of data observed. Analyses not included in the paper revealed that the Harris Model might be able to account for the results, too (see Conclusions section for a detailed discussion.) Furthermore, the convergent results in both eyelid conditioning and contingency learning confirm the generality of this finding and of the rules that govern different forms of human learning. Divergent results observed in other studies may be due to additional differences in their experimental procedures concerning stimuli and stimulus presentation.

2.2. Generalisation decrements

One of the most relevant advantages of both modern elemental models compared to the Rescorla-Wagner Theory is the ability to predict and explain external inhibition. External inhibition describes the decrement in responding to a successfully trained stimulus if a new unknown stimulus is added, i.e. if the previous trained stimulus is presented within a new context. Because of the context-independent representation of the trained stimulus in the Rescorla-Wagner Theory, the prediction and the resultant response are context-independent, too, and no decrement is expected. In contrast, the context-dependent representation in both the Replaced Elements Model and the Harris Model leads to a generalisation decrement if a stimulus is presented in a different context. However, both models are not flexible in their predictions concerning the symmetry of generalisation decrements due to adding components to a previously trained stimulus compared to removing a component. Instead, both theories predict asymmetrical generalisation decrements. More precisely, adding a component B after training A+, i.e. testing AB?, always produces less of a decrement than removing the component B after training AB+, resulting in test trial A?.

In the Replaced Elements Model, when something is added to a stimulus configuration, its context-dependent elements are replaced. When something is removed, context-dependent elements of the remaining components are also replaced and, in addition, all elements of the removed component are no more activated. Returning to our example, in the "added cue" task, a_i and $a_{\sim b}$ will gain associative strength during training with A+. In the "removed cue" task, a_i, b_i, a_b, and b_a will compete for the associative strength while AB+ is trained. In the "added cue" test (AB?), a_i, b_i, a_b, and b_a will be activated. Therefore, only the associative strength of a_i will remain active to elicit a response. This is true for the "removed cue" task, too, as a_i and the untrained $a_{\sim b}$ will be activated in test trials (A?). But in the "added cue" task, a_i will have gained twice as much associative strength compared to the "removed cue" task as there are fewer elements to compete against during training. This is independent of the value of *r*. (There is an exception in the case of $r = 0$. In this case, the decrements are symmetrical because no generalisation at all is predicted. But this

case is primarily of theoretical importance as the assumption of "no generalisation at all" is in conflict with empirical observations.)

In the Harris Model, the generalisation decrement in the “removed cue” task is determined by three processes. First, the distinct elements of the removed component B are no more activated at all in the A? trials. Second, the common elements of A and B are excited by A alone and their activation decreases. Third, elements of A that were displaced from the buffer in AB+ training trials gain entry into the buffer in the A? trials and their activation is boosted. The net effect of these opposite processes is quite stable and independent on the proportion of common elements. In AB? test trials of the “added cue” task, some previously trained distinct elements of A are displaced from the buffer and their contribution to the associative strength is reduced. The resulting decrement is however always smaller than in the “removed cue” task where the whole associative strength gained by the elements of B is lost. Furthermore, the activation of all elements common to A and B is doubled due to the summed input and some of these common elements are then additionally boosted by the buffer. Hence, if there are enough common elements, the summed activation of elements of A is higher in compound AB? trials than in single stimulus A+ trials and instead of a decrement, an increment in responding is predicted. Simulations with other values for gain and power confirmed this conclusion. Therefore, the Harris model always predicts an asymmetric generalisation decrement independent of the values of these parameters.

The Pearce Model as well as its extension predict symmetrical decrements. The decrements in the Pearce Model are due to co-activation of similar configural units in the generalisation test, and as the similarity is symmetrical, generalisation decrements are symmetrical, too. More precisely, as the similarity between A and AB is .5 for $d = 2$, the response to the test configuration is predicted to be half the size of that to the trained configuration, irrespectively of which one of the two stimulus configurations is trained and which is tested. Even if one varies the discrimination parameter d, the relative size of generalisation decrements in the two generalisation tests remains symmetrical.

Hence, comparing generalisation decrements after adding and after removing components allows a comparison of the predictive power of the configural model of Pearce (1994) against the elemental models of Wagner (2003) and Harris (2006).

2.2.1 Symmetrical generalisation decrements: Configural stimulus processing in human contingency learning.

In previous conditioning and causal learning studies, the empirical decrements were always asymmetrical. This should allow to investigate whether there are stimulus properties that can trigger a shift from an elemental to a configural processing mode. In three contingency learning experiments, coloured dots served as predictors for the temperature of the cooling water in a nuclear power plant. Furthermore, we manipulated the motion and the spatial arrangement of these dots in order to induce configural or elemental processing by perceptual grouping.

In Experiment 1, participants were trained with compounds of either two or five dots that were moving across the black computer screen for 2 s. In the test stage, both five- and two-dots compounds were presented to all participants. Thus, for some participants three dots were removed in the test stimulus; for the remaining three dots were added. An additional factor controlled the direction of the movements. In one condition, all dots were moving in the same direction, randomly selected for each trial and therefore irrelevant for predictive learning; in another condition, a direction was randomly selected for each dot separately. The question was whether the grouped motion would be sufficient to induce configural processing and therefore symmetrical decrements. The decrements, however, were already symmetrical in the “separated motion” condition.

In order to investigate whether motion per se caused the symmetrical decrements in Experiment 1, stationary dots were used in Experiment 2 and 3. In Experiment 2, the dots were presented on variable positions randomly selected across the entire screen. Furthermore, the number of dots was reduced from five to three in the training compound of Group Removed Cue and the test compound of Group Added Cue. Again, symmetrical decrements were observed. This was replicated in one experimental group

of Experiment 3, whereas an additional group received the dots on fixed positions evenly distributed in the middle of the screen. (Generalisation task was now manipulated within participants.) The results reliably showed symmetrical decrements for added and removed components. The manipulations of the stimuli arrangement had no effect on the decrements.

All results are in line with the configural model of Pearce, but at variance with the elemental models as well as with results of previous studies. Three explanations for the symmetrical decrements in the present experiments seem reasonable: Time pressure during stimulus processing could have resulted in participants choosing configural strategies; the non-causal cover story could have reduced cue competition effects in the "removed cue" task and thus its decrement; and finally, generalisation not only between the stimulus compounds but also between each component, i.e. each dot, could have produced symmetrical decrements based on elemental processing. The first two of these points were addressed in two additional experiments (see below).

2.2.2 Generalization decrements and salience: new data and more elaborated analyses

In two experiments, the effects of time pressure during stimulus processing and the influence of cover story on generalisation were investigated. In the first experiment, time pressure was controlled as a between-subjects factor: coloured dots were presented for either 2 or 8 s. Additionally, a causal cover story was introduced (bacteria causing illness) and manipulations similar to Experiment 1 of chapter 2.2.1 were used in order to investigate again the effect of perceptual grouping: three or two dots were moving across the screen either in the same direction, or each dot moved in a direction randomly chosen. Surprisingly, asymmetrical decrements were obtained independently of the manipulations introduced. In all groups, removing an already trained component from a compound resulted in larger decrements than adding a new one. This is at variance with the experiments of chapter 2.2.1 that used the same stimulus material and arrangement but resulted in symmetrical decrements. Thus, neither perceptual grouping by motion nor time pressure is sufficient to cause symmetrical decrements.

However, the cover story could have effectively controlled the decrements and stimulus processing as in the experiments of Chapter 2.2.1 a predictive cover story was used. This hypothesis was tested in the second experiment. Whereas one group received the causal instruction of Experiment 1 (bacteria causing illness), a second group was told that they should learn something about blood components that indicate an illness. Again, asymmetrical decrements were observed in both groups. As introducing the predictive cover story did not result in symmetrical decrements, the causal cover story could not have been the sole reason for the asymmetry in Experiment 1 and the resulting divergence to the experiments of the chapter 2.2.1.

In order to further look into this divergence, more precise analyses based on single test trials were carried out. Each training stimulus consisted of at least two different coloured dots and each of them was sufficient to solve the discrimination. Therefore, participants may have learned differently about each coloured dot and the colours within compounds may have been perceived as differently relevant by different participants. This becomes crucial in the test of the “removed cue” task where in each of three test trials another, and therefore presumably differently relevant, dot was removed. This should result in different rating for the different "removed cue" test compounds. In order to examine this in more detail, we identified the minimal and maximal ratings of each participant for all trial types. This analysis exposed a new grouping of the experiments. Before, when ratings averaged over all test trials were analysed, we contrasted the experiment of chapter 2.2.1 with the present experiments based on their symmetrical versus asymmetrical decrements. The underlying pattern, which resulted in these averaged decrements, revealed that, while the decrements were symmetrical in both minimal and maximal rating in Experiment 2 and 3 of chapter 2.2.1, an asymmetry in favour of the removed component increased from the maximal to the minimal rating in the current experiments as well as in Experiment 1 of chapter 2.2.1. Actually, a switched asymmetry in the maximal ratings with a smaller decrement for the removed component was observed in Experiment 1. In terms of modelling, the perceived relevance of a dot can be implemented by assigning different saliences to different relevant dots. Additional simulations revealed that elemental models failed to

explain the symmetrical decrements in the minimal and maximal ratings, even if one specifies different saliences. In contrast, all data were in line with the configural model of Pearce (1994) when different saliences for each dot are assumed for the current experiments and Experiment 1 of chapter 2.2.1 but same saliences in the case of Experiment 2 (and 3) of chapter 2.2.1. Thus, the response pattern in the minimal and maximal ratings and the results of the simulations suggests that the configural Pearce Model can predict asymmetrical decrements and that elemental models are not necessary to explain the observed asymmetry. Additionally, we suppose that different saliences within the compounds become only relevant if more and more complex stimuli have to be processed.

Conclusions

The present thesis demonstrated that models that comprise configural principles are not made superfluous by modern elemental models. Using computer simulations, it was demonstrated that the models make contradicting predictions, despite their flexibility. Testing these predictions, neither the Replaced Elements Model nor the Harris Model was confirmed in all experiments. Instead, the configural model of Pearce (1994) was able to account for all behaviour observed both during learning (feature-negative discrimination) and after learning (generalisation decrements) as well as in different paradigms (eyelid conditioning, causal and predictive learning). However, in case of the acquisition of feature-negative problems this is only true if also some flexibility is introduced to the configural model as suggested by Kinder and Lachnit (2003) in the extended Pearce Model. Together with the “elemental” results of other studies, i.e. results that can so far only be explained by elemental models, this corroborates the position that there are two modes of stimulus processing during associative learning. Further tests and experiments however are necessary to investigate factors that influence stimulus processing and representation, the second aim of the thesis. The results revealed that their effects are not inevitable. In two different learning paradigms, identical results were obtained in the feature-negative discriminations. Perceptual grouping by motion was also no strong manipulation in the

generalisation experiments. In none of them, it had a consistent influence. Furthermore, the influence of predictive versus causal cover story, which is reported in the literature, was also not visible in my data. Instead, the number and complexity of the stimuli seemed to have influenced the perceived relevance of the stimuli in the generalisation task but not their representation: the relative size of saliences differed between the experiments but all data were in line with the configural representation proposed by Pearce. This indicates that more controlled research is needed to investigate not only the factors but also the circumstances controlling these factors' relevance.

In conclusion and considering the theoretical implications of my thesis, what are the reasons for the failure of the elemental models?

As described, the predictions of the Replaced Elements Model for the feature-negative learning task originate in the assumed higher overall activation due to an added component. The context-dependency within the model is only qualitative. The *r* parameter controls which elements are activated but not how many. Instead, the amount of activation for a stimulus compound is still the linear sum of the activation provoked by its components. Therefore, learning will be faster the more components a configuration has, independently of the kind of components, their perceptual interaction or, as in the feature-negative study, their relevance for the learning problem. This is also the reason why the Replaced Elements Model cannot predict symmetrical generalisation decrements – or even "switched" asymmetries, i.e. a smaller decrement after removing than after adding something, in test trials in which the least relevant component is removed. The higher number of elements activated in the "removed cue" compared to the "added cue" training compound ((a_i, b_i, a_b, b_a) versus (a_i, $a_{\sim b}$)) results in more cue competition in the first than in the latter training and thus in relatively less associative strength for each element in the "removed cue" compound. Even if the salience of the to-be-removed elements b_i and b_a is zero and they do not compete for associative weight during training, a_i (i.e. the elements that remain active in both test compounds A? and AB?) cannot have gained more associative strength in the "removed cue" than in the "added cue" training compound.

Thus, the response cannot be higher and the decrement not smaller in the first than in the latter test compound.

In the Harris Model, despite the limited capacity of the attention buffer, there is also no upper limit for the elemental representation. Instead, it is still true that the more stimuli are around, the more elements will be activated and the faster learning happens. However, the effect of more activation by more presented stimuli can be counteracted by the attention buffer. The capacity of the buffer controls the context-dependency of a representation. Normally, it is set to contain exactly or little less than all elements of one stimulus. For the feature-negative task, this induces maximal context-dependency for the ordinary feature-negative problem. All elements of A will be in the buffer during A+ trials, but only half of A's activation weights will stay there when presented in AB- trials. In the CD+ trials, there are already some elements of C and D outside the buffer and these elements will not change their activation in CDE-trials. Therefore, the representation of C and D is less context-dependent than that of A and this more similar pattern of activation results in relatively more generalisation between CD+ and CDE- than between A+ and AB-. Thus, the Harris Model can predict, for the same reason as the Pearce Model, that adding a common cue to a discrimination impedes its learning. This effect disappears immediately when the buffer's capacity is lower or higher. Then, more elements are always, i.e. context-independent, outside or inside the buffer, respectively, and the relative difference in the context-dependency ultimately vanishes. Instead, the effect of more activation in the CD+/CDE- discriminations, that determines the predictions of the Replaced Elements Model and the Rescorla-Wagner Theory, also controls the prediction of the Harris Model and adding a common cue is predicted to facilitate discrimination learning. With respect to the generalisation decrements, the Harris Model is able to predict a "switch" of the asymmetry when a less salient component is removed. Simulations in parallel to that of the Pearce Model, however, showed that the decrements averaged over different saliences will always be larger in the "removed cue" than in the "added cue" task. The stronger cue competition in the removed task described for the Replaced Elements Model dominates the predictions.

In contrast, the Pearce Model can deal with the present data because the activation strength of the configural unit that represents the currently presented stimuli and whose association will change is fixed and independent of the complexity of the stimulus configuration or the intensity of the stimulus components. This is because the activation of the input vector is normalised, and therefore fixed, as implied in Equation (1). Concluding from our results, future associative learning focussing on stimulus processing and representation models may include more elaborated assumptions on the general processing capacity. Interestingly, Wagner and Brandon (2001, also see Brandon, Vogel & Wagner, 2000) on their way to the Replaced Elements Model developed an elemental model that incorporates Pearce fixed activation assumption. In their so called "inhibited elements model", the rule is that if A is coactive with B, A inhibits half of the elements otherwise activated by B and vice versa. Thus, the total number of representational elements in any compound is a constant. In the Pearce Model, not the number of elements is fixed but the activation that is distributed over the elements. However, saying that there is always an overall activation of 1 or that there are always 100 elements activated is the same, as long as the activation of the elements is fixed to .01 and the salience of a component is not conceptualised as the amount of activation (as in the Harris Model) but is defined by the number of elements activated by the component. With respect to our results, it seems that elemental models have to go further this way. Allowing flexibility in the proportion of inhibited elements should result in a model that comprises the Pearce Model at one end, the Harris Model in the middle (there, elements are not totally inhibited but for some, the activation is reduced) and the Rescorla-Wagner Theory at the other end of its spectrum. As Figure 1 indicates for generalisation decrements, this elemental model is able to deal with symmetrical decrement observed in this thesis but also with the asymmetry in previous studies. Furthermore, it should have no problems with summation and other cue competition effects where the Pearce Model still gets into trouble.

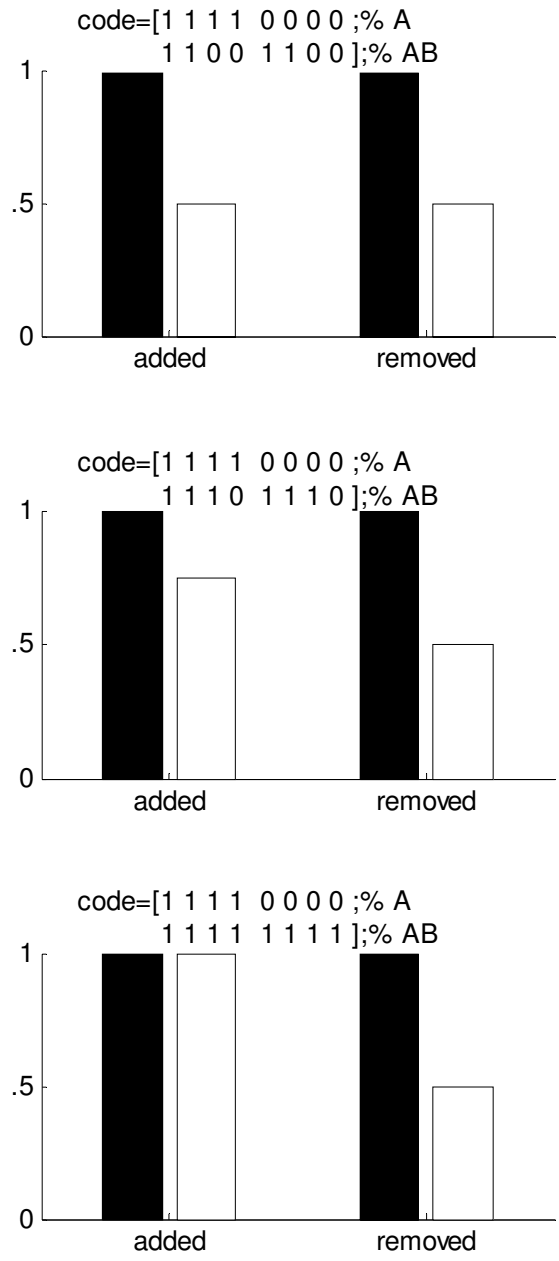


Figure 1. Predictions of an inhibited elements model assuming different amounts of inhibited elements for the generalisation tasks. In the top panel, half of A's elements are inhibited when presented in compound AB, resulting in a constant number of four activated elements as proposed by Wagner and Brandon (2001) and symmetrical decrements. In the middle panel, simulations revealed asymmetrical decrements if only one of four elements is inhibited. In the bottom panel, nothing is inhibited, matching the Rescorla-Wagner Theory, and the predictions show an asymmetrical decrement with missing external inhibition in the "Added Cue" task.

References

Bouton, M., Mineka, S. & Barlow, D. (2001). A modern learning theory perspective on the etiology of panic disorder. *Psychological review*, *108*, 4-32.

Brandon, S., Vogel, E. & Wagner, A. (2000). A componential view of configural cues in generalization and discrimination in Pavlovian conditioning. *Behavioural Brain Research*, *110*, 67-72.

Dickinson, A. (2001). The 28th Bartlett Memorial Lecture - Causal learning: An associative analysis. *Quarterly Journal of Experimental Psychology, 54B*, 3-25.

Glautier, S. (2002). Spatial separation of target and competitor cues enhances blocking of human causality judgements. *Quarterly Journal of Experimental Psychology*, *55B*, 121-135.

Glautier, S. (2008). The elemental-configural distinction: A problem of two dimensions - Commentary on Melchers, Shanks, and Lachnit (2007). *Behavioural Processes*, *77*, 431-433.

Gluck, M. A. & Bower, G. H. (1988). From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General*, *117*, 227-247.

Harris, J. A. (2006). Elemental representations of stimuli in associative learning. *Psychological Review*, *113*, 584-605.

Haselgrove, M., Robinson, J., Nelson, A. & Pearce, J. M. (2008). Analysis of an ambiguous-feature discrimination. *Quarterly Journal of Experimental Psychology*, *61*, 1710 - 1725.

Kinder, A. & Lachnit, H. (2003). Similarity and discrimination in human Pavlovian conditioning. *Psychophysiology*, *40*, 226-234.

Lachnit, H. (1988). Convergent validation of information processing constructs with Pavlovian methodology. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 143-153.

Lachnit, H., Schultheis, H., König, S., Üngör, M. & Melchers, K. (2008). Comparing elemental and configural associative theories in human causal learning: A case for attention. *Journal of Experimental Psychology: Animal Behavior Processes*, *34*, 303 - 313.

Livesey, E. J. & Boakes, R. A. (2004). Outcome additivity, elemental processing and blocking in human causality judgements. *Quarterly Journal of Experimental Psychology*, *57B*, 361-379.

Livesey, E. J. & Harris, J. A. (2008). What are flexible representations? Commentary on Melchers, Shanks and Lachnit. *Behavioural Processes*, *77*, 437-439.

Mackintosh, N. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review*, *82*, 276-298.

Melchers, K. G., Lachnit, H. & Shanks, D. R. (2004). Past experience influences the processing of stimulus compounds in human Pavlovian conditioning. *Learning and Motivation*, *35*, 167-188.

Melchers, K. G., Shanks, D. R. & Lachnit, H. (2008). Stimulus coding in human associative learning: Flexible representations of parts and wholes. *Behavioural Processes*, *77*, 413-427.

Miller, R. R. & Matzel, L. D. (1988). The comparator hypothesis: A response rule for the expression of associations. In G. H. Bower (Ed.), *The psychology of learning and motivation* (*vol. 22*, pp. 51–92). San Diego, CA: Academic Press

Miller, R. R., Barnet, R. C. & Grahame, N. J. (1995). Assessment of the Rescorla-Wagner Model. *Psychological Bulletin*, *117*, 363-386.

Myers, K., Vogel, E., Shin, J. & Wagner, A. (2001). A comparison of the Rescorla-Wagner and Pearce models in a negative patterning and a summation problem. *Animal Learning & Behavior*, *29*, 36-45.

Pearce, J. M. (1987). A model for stimulus generalization for Pavlovian conditioning. *Psychological Review*, *94*, 61-73.

Pearce, J. M. (1994). Similarity and discrimination: A selective review and a connectionist model. *Psychological Review*, *101*, 587-607.

Pearce, J. M. (2002). Evaluation and development of a connectionist theory of configural learning. *Animal Learning & Behavior*, *30*, 73-95.

Pearce, J. M. (2009). The 36th Sir Frederick Bartlett lecture: an associative analysis of spatial learning. *Quarterly Journal of Experimental Psychology, 62*, 1665-1684.

Pearce, J. M. & Bouton, M. E. (2001). Theories of associative learning in animals. *Annual Review of Psychology*, *52*, 111-139.

Pearce, J. M., Esber, G. R., George, D. N. & Haselgrove, M. (2008). The nature of discrimination learning in pigeons. *Learning & Behavior*, *36*, 188-199

Pearce, J. M. & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, *87*, 532-552

Pearce, J. M. & Redhead, E. (1993). The Influence of an Irrelevant Stimulus on Two Discriminations. *Journal of Experimental Psychology: Animal Behavior Processes*, *19*, 180-190.

Rescorla, R. & Wagner, A. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. Black & W. Prokasy (Eds.), *Classical conditioning II: Current research and theory (pp* 64–99). New York: Appleton-Century-Crofts

Wagner, A. (1981). SOP: a model of automatic memory processing in animal behavior. In N. Spear & R. Miller (Eds.), *Information Processing in Animals: Memory Mechanisms* (pp. 5–47). Hillsdale, NJ: Erlbaum.

Wagner, A. R. (2003). Context-sensitive elemental theory. *Quarterly Journal of Experimental Psychology*, *56B*, 7–29.

Wagner, A. R. (2008). Evolution of an elemental theory of Pavlovian conditioning. *Learning & Behavior*, *36*, 253-265.

Wagner, A. R. & Brandon, S. E. (2001). A componential theory of Pavlovian conditioning. In R. Mowrer & S. Klein (Eds.), *Handbook of Contemporary Learning Theories* (pp. 23-64). Mahwah,NJ: Erlbaum.

Waldmann, M. R. (2007). Combining versus analyzing multiple causes: How domain assumptions and task context affect integration rules. *Cognitive Science*, *31*, 233-256.

Walther, E., Nagengast, B. & Trasselli, C. (2005). Evaluative conditioning in social psychology: Facts and speculations. *Cognition & Emotion*, *19*, 175 - 196.

Whitlow, J. & Wagner, A. (1972). Negative patterning in classical conditioning: Summation of response tendencies to isolable and configural components. *Psychonomic Science*, *27*, 299-301.

Young, M. & Wasserman, E. (2002). Limited attention and cue order consistency affect predictive learning: A test of similarity measures. *Journal of Experimental Psychology: Learning*, *Memory*, *and Cognition*, *28*, 484-496.

Reprints and Manuscripts

Behavior Research Methods
2009, 41 (1), 29-34
doi:10.3758/BRM.41.1.29

ALTSim: A MATLAB simulator for current associative learning theories

Anna Thorwart
Philipps-Universität Marburg, Marburg, Germany

Holger Schultheis
Universität Bremen, Bremen, Germany

and

Stephan König and Harald Lachnit
Philipps-Universität Marburg, Marburg, Germany

ALTSim is a MATLAB-based simulator of several associative learning models, including Pearce's configural model, the extended configural model, the Rescorla–Wagner model, the unique cue hypothesis, the modified unique cue hypothesis, the replaced elements model, and Harris's elemental model. It allows for specifying all relevant parameters, as well as exact stimulus sequences by graphical user interfaces. It is an easy-to-use tool that facilitates evaluating and comparing the featured associative learning models. ALTSim is available free of charge from www.staff.uni-marburg.de/~lachnit/ALTSim/.

During the past few decades, a fair amount of research in the field of associative learning was directed toward the development of formal models. The explanatory and predictive power of such models has to be tested in experiments designed to allow for comparisons between models (e.g., Lachnit, Schultheis, König, Üngör, & Melchers, 2008). In this context, it is a significant advantage of most current associative learning models that they employ a mathematical model that allows users to derive detailed predictions. With earlier models (like the Rescorla–Wagner model; Rescorla & Wagner, 1972), these predictions mostly could be accomplished by relatively simple means, such as by paper and pen. As models became more complex and sophisticated over the years, computer programs became necessary (see, e.g., Glautier, 2007; Schultheis, Thorwart, & Lachnit, 2008a, 2008b). The present simulator, called ALTSim, follows this tradition but goes beyond it; it includes several models in a common program environment. Because of this, it further facilitates direct comparison of the models included and allows the generation of meaningful designs and discrimination tasks.

The following models are included in the current version of ALTSim and will be described below: Pearce's configural model, the extended configural model, the Rescorla–Wagner model, the unique cue hypothesis, the modified unique cue hypothesis, the replaced elements model, and Harris's elemental model. Whereas all these models assume error-based learning rules similar to the one in the Rescorla–Wagner model, they differ in aspects concerning the coding, representation, and processing of stimuli. In configural models, a stimulus compound is represented and associated as a whole, independently of its components. Elemental models assume that individual components of a stimulus compound are associated with an outcome. For a more detailed discussion, see Melchers, Shanks, & Lachnit, 2008, with commentaries.

The models differ in their nomenclature, so the following apply specifically to the present article and simulator. A conditioned stimulus (CS) is any stimulus or stimulus compound presented in any contingency with the unconditioned stimulus (US)—that is, it does not imply that a CS consists only of a single (distinct) stimulus. The presence of *stimulus compound* or *stimulus configuration* indicates that a CS comprises several individual stimuli, which are referred to as *components*. *Element* describes the hypothetical representational unit assumed by some elemental models. *Salience* describes a property of a CS that controls the amount of learning (learning rate parameter α).

In two articles, Pearce (1987, 1994) presented his configural model. The version included in ALTSim reflects the configural model of Pearce (1994). It assumes that for each CS, independent of the number of its components, a single configural unit is created and associated with the US. Learning (i.e., change in the associative strength) occurs only between the configural unit of the CS presented and the US. However, other configural units are activated, depending on their similarity with the presented CS, and in this way contribute to the observed conditioned re-

A. Thorwart, anna.thorwart@staff.uni-marburg.de

sponse (CR). The similarity between two CSs is described by the following equation, with n_c = number of common components to both CSs, n_X = number of components of CS X, and n_Y = number of components of CS Y:

$$S = \left(\frac{n_c}{\sqrt{n_X * n_Y}} \right)^2 .$$

Kinder and Lachnit (2003) extended the model by assuming a free generalization parameter d (substituting d for 2 in the exponent) that allows one to control the overall amount of discriminability between stimuli (extended configural model); hence, with $d = 2$, the similarity computation in the extended model corresponds to Pearce's configural model. With increasing d, the generalization gradient becomes steeper.

The Rescorla–Wagner model (Rescorla & Wagner, 1972) is also included in ALTSim. It assumes that a CS is split into its components during associative learning; that is, each component is represented by a single distinct element and associated on its own. However, it is the associative strength of the entire CS that is used in the error term of the learning rule of each component. The associative strength of a compound equals the sum of the associative strengths of its components. Therefore, the Rescorla–Wagner model is a linear elemental model, in which learning about a component depends on all components present in a trial.

The Rescorla–Wagner model is very successful. However, due to its linearity, it is unable to predict successful learning in nonlinear discrimination problems, such as negative patterning. In order to circumvent this problem, an extension called the unique cue hypothesis was suggested (Whitlow & Wagner, 1972). Here it is assumed that a CS compound is different from the mere sum of its components. Instead, when components are presented together in a compound, it is hypothesized that additional unique cues that represent unique properties of combinations of components emerge. Unique cues are functionally equivalent to the other components and enter into their own association with the US. When one includes unique cues, the Rescorla–Wagner model becomes a nonlinear elemental model, because a compound activates elements other than just its own components. Redhead and Pearce (1995) argued that the rate of learning about a component may be reduced when it is accompanied by other components. Therefore, they introduced a modified salience computation within the unique cue hypothesis, whereby conditioning of a component is influenced by the number and saliences of the stimuli that accompany it. For further details, see Equation 2 of Redhead and Pearce.

Wagner's replaced elements model (REM; 2003) is also a nonlinear elemental model. Each stimulus component is represented by a number of elements that are associated separately with the US, following the Rescorla–Wagner learning rule. Some of these elements are context independent; that is, they are always activated when their component is present. Others are context dependent; that is, there are elements that are activated only when their component is presented without another component, and there are elements that are activated only when their component is present together with the other component. Thus, the latter replace the former whenever a component is presented in compound. The proportion of elements that are replaced in a component's representation is conceptualized as a free parameter called the replacement parameter r. Glautier (2007) and Schultheis et al. (2008b) developed two differing mathematical implementations of the replacement. Whereas r constitutes the relative activation of the different CS parts in Schultheis et al. (2008b), it is a learning parameter for Glautier and is functionally similar to alpha. Therefore, the associative weights differ between the two implementations as $V_{\text{Glautier}} = r * V_{\text{Schultheis et al.}}$. The output activation, however, is the same, with the exception of the first trial after changing r.

Harris's elemental model (2006) quantitatively assumes nonlinearity. As in the REM, each component is represented by several elements that are associated individually. A compound is represented by the same elements that are activated by its individual components. However, any element's activation by a compound differs from its activation by a lone component. Elements compete for attention on the basis of their salience, and any element that captures attention receives a further boost of its activation. Because of this competition among elements, some elements that are boosted when their components are presented alone are activated weakly when their components are presented in compound. Because the activation strength serves as salience parameter in the learning rule, learning about these elements differs depending on other components' being present in a trial.

ALTSIM'S USER INTERFACES

ALTSim has several graphic user interfaces (GUIs). In the main GUI, parameters and configurations needed for all models are specified, and the model to be simulated is chosen. Once this choice is made, there are additional GUIs in which specific parameters are defined for the chosen model. All parameters must be specified unless explicitly mentioned in the description below. Default values, however, are predefined in a separate file called "default.conf" and are saved in the same directory as the simulator.

Main GUI

The main GUI is started by running ALTSim.exe or ALTSim.m within MATLAB. It comprises three panels: stimuli, stimulus sequence, and model. In the stimuli part, two fields allow the user to specify the full path of files relevant to the simulation. For each field, the path is specified either by hand or by using the "Search" button on the right to pick the desired file. Both files are simple text files with values separated by white space. The files can be examined and changed by clicking on the "View" button. Note that this only works for MATLAB versions

2008a or higher. Furthermore, within ALTSim, one can change only existing files and cannot expand such files with new stimuli or trials. To add new stimuli or trials, one must change the file outside ALTSim. We hope this will be solved in a future version of MATLAB.

The first field, *trial matrix*, holds the path to the file containing definitions of each trial type. In the trial matrix, each line consists of three entries that describe one trial. The first, a string, is the name of the trial used in the results graphs and can be chosen freely. The remaining two entries are numbers specifying the learning parameters lambda and beta of the US. For the Harris model, the value for beta is divided by 10 in order to avoid positive feedback. As is the convention, the values should range from 0 to 1. Additionally, for all but the Harris model, it is possible to assume negative values for lambda. Since the activation of the US elements in the Harris model is not yet finally conceptualized, trials with negative lambda values will be treated as nonreinforced trials, with lambda = 0. In the case of test trials (i.e., trial types that are not specified in the training sequence), the values of lambda and beta are ignored during evaluation.

The second field, *stimulus code*, holds the path to the file containing the definitions for the CS in each trial. One line in the stimulus code specifies the CS components to be presented in the corresponding trial of the trial matrix. Each column represents one of the CS components: "1" indicates that this CS component is present, and "0" indicates that the CS component is absent. All trials—training *and* test—must be defined in the trial matrix, as well as in the stimulus code.

In the *stimulus sequence* panel, the sequence of trials that should be simulated is specified. The panel has two modes: It is possible to generate new stimulus sequences within ALTSim or to load an existing stimulus sequence. The drop-down menu on the right side of the panel allows one to select either of these two modes. After selection, the appropriate fields become visible. In both cases, the fields *number of sequences* and *training block* must be specified. Number of sequences specifies the number of different stimulus sequences or participants in an experiment. The simulation is computed for each sequence and then averaged across all sequences to get the results. In the training block, all trials of the trial matrix that are presented (trained) during the simulation must be specified. They must be named by their line numbers in the trial matrix and separated by white space.

In the case of loading sequences, there is an additional field, *load sequence*, for the path to the sequence file. One can specify the path either by hand or by using the "Search" button on the right to find and select the desired file. In the sequence file, one stimulus sequence corresponds to a column of numbers referring to the appropriate trial of the trial matrix. The columns should be separated by white spaces or commas.

A sequence is generated by replicating the training block. To manipulate the relative frequency of one trial type, it can be specified several times within the training block field. The number of repetitions of the entire training block is specified in the *numbers of blocks* field. Thus, simulating a training block of 7 trials and 100 blocks will yield 7 * 100 presentations of a CS and corresponding weight updates. The "randomize" check box enables randomization of the sequence order. In this case, the maximal number of subsequent trials with the same lambda must be specified as a constraint on the generation of the randomized order of sequences. If a file path is specified in the *save sequence* field, the generated sequences are saved in a .txt file. This field is optional.

Now the "1. initialize stimuli" button can be clicked to prepare the stimuli, parameters, and sequences for the simulation. By clicking this button, one can check whether all necessary information has been specified. If it has, all relevant files and fields will be opened and read, and the stimuli and sequences will be set up. These settings are used for all simulations until the button is clicked again or the main GUI is closed.

In the theory panel, a list box specifies all available models. By clicking the "2. choose model" button, one causes the parameters to be passed to the GUI of the selected model, which will be opened in front of the main GUI.

Specific GUIs

Common characteristics. Even though the GUIs are specific to each model, they are arranged in a similar way and comprise some common fields. These include the *starting weights* and *alpha* fields. In the *starting weights* field, one can specify the path to a file containing associative weights that were learned in a previous simulation. One can specify the path either by hand or by using the "Search" button on the right to find and select the desired file. If the field is empty, the starting values are zeros. If a valid path is read in, the values in the last line of the file are used as starting values of the associative weights for the simulation. This permits the simulation of several "phases" with different parameters or stimuli within one "experiment"; for example, in the REM, one can train with $r = .6$, then change r and simulate again, this time with the previously trained associative weights as the starting point. It is therefore necessary that both simulations include the same numbers of CS components. Due to its specifics, the Harris model simulator does not include this functionality. In order to simulate several phases with different stimuli, however, an appropriate stimulus sequence can be loaded. As an alternative, Schultheis et al.'s (2008a) specialized Harris simulator can be used.

In the alpha field, the salience parameter is set. The values are separated by white spaces. Even if all models assume an alpha parameter as a characteristic of the CS, they conceptualize it in different ways. Therefore, the number of alphas needed depends on the selected model. As a guide, the correct number of alphas is indicated in a question dialog box appearing if the user starts the simulation with a wrong number of alphas (see below).

In the lower part of each GUI, one can indicate what should be done with the simulation results. In the *save results in* field, one can specify a place and a prefix for the two separate text files that the averaged associative weights and response activations are saved in; the names of the written-out files then consist of the user-defined prefix and the extensions "_weights.txt" and "_testactivion.txt," respectively. In both files, each line displays the results after one trial. Each column contains the results for a hypothetical CS unit or a trial type, respectively. Furthermore, one can selectively disable and enable graphical output using the "plot associative weights" and "plot output activation" check boxes. In the associative weights graph, the associative weights of the hypothetical CS units are displayed. The response to all CS compounds defined in the trial matrix at the beginning of each trial is shown in the activation by test stimuli graph. Note that in both graphs' titles, alphas and betas are not displayed if there are more than seven. As indicated before, the Harris model differs from the other models in this respect. For conceptual reasons, only the output activations are plotted, but no weights are saved or plotted.

The simulation is started when the user clicks the "3. simulate" button. The remaining parameters are read in, and an appropriate stimulus representation is created. If the number of alphas does not fit with the stimulus representation, a question dialog box appears. There, one can choose either to use the default value for all alphas or to cancel the simulation and enter the correct alphas manually. Learning is then simulated on a trial-by-trial basis and, finally, the results are saved and plotted as indicated.

Pearce's configural model. In addition to the common fields, the generalization parameter *d* has to be specified for simulating Pearce's configural model. The value should be a positive number. Furthermore, for each trial type in the trial matrix, there must be an alpha defined, even if it is the same CS (i.e., the same configural unit), but with different lambdas or betas.

Rescorla–Wagner model (with modifications). The Rescorla–Wagner GUI is selected in order to simulate the original Rescorla–Wagner model, the unique cue hypothesis, or Redhead and Pearce's (1995) modification. These alternatives are enabled and disabled by using two additional check boxes in the central part of the GUI. If the check box "compute unique cue" is checked, unique cues are assumed for all possible combinations of the CS components; for example, for three CS components, such as A, B, and C, the following unique cues are generated: AB, AC, BC, and ABC. This example also indicates the general order of the unique cues. The unique cues are created by permutation of all combinations, beginning with the two-component CS compounds containing the first component, then those containing the second component, and so on. Afterward, all combinations with three components are generated, beginning with combinations containing the first two components, then the first and the third component, and so on. The check box "use modified salience equation of Redhead and Pearce (1995)" controls whether the original alpha or the modified equation proposed by Redhead and Pearce is used. In any case, the Rescorla–Wagner model assumes one alpha parameter for each CS component and for each unique cue.

Replaced elements model. The *proportions matrix* field specifies the full path, which can be either explicitly typed in or chosen using the "Search" button, of the file containing the proportions of replaced (r) and not replaced ($1 - r$) elements of any CS component, in the context of any other CS component. Proportion files have the file extension ".pro." The proportions matrix contains n lines, each representing proportions for one of n CS components. Each of these lines again contains $n - 1$ real numbers. For any line L, these numbers indicate the proportions of elements to be replaced for CS_L in the context of the other CSs. To put it more precisely, r_{ij} (with i the index of the line and j the index of the column) is the proportion of elements of CS_i that will be replaced in the presence of CS_j. Since the numbers in each line represent proportions of to-be-replaced elements, they should lie within the closed interval [0,1]. The file can be examined and changed by clicking the "View" button. Since Schultheis et al. (2008b) and Glautier (2007) developed two differing mathematical implementations of replacement, the list box "learning rule" controls which should be used for simulations. Furthermore, there is only a general alpha parameter controlling the overall learning rate in the REM.

Harris's elemental model. A number of parameters must be specified for the Harris model. Since the model uses random numbers to determine the intensities of the elements of each stimulus and the existence of connections among elements, different runs of the model may give different results. To get an impression of how the model will behave on average, the model is simulated x times, and the average result of these x runs is taken as the model output. If there is more than one sequence to simulate, all sequences will be simulated within a run and then averaged. The "Number of runs" parameter indicates x. A free parameter of the model, "Elements per CS," determines the number of elements sampled for each CS. This parameter can be specified by using the according field. Harris's model assumes that the involved elements are only partially connected (i.e., that there exist certain pairs of elements that have no connection to each other). By changing the "connection density" parameter, one can control for the amount of connections existing between pairs of elements. Valid entries for this parameter are all real numbers from the closed interval [0,1], where 0 results in no connections at all and 1 results in complete connection of all elements. The "gain" parameter allows one to specify the amount of boost resulting from the attention buffer. To put it more precisely, the increase of activation of each element inside the buffer will be multiplied by the gain parameter value. Furthermore, for each CS component, an alpha value must be specified.

There are essentially two ways of specifying the capacity of the attention buffer. One option is to have the simulator set up the capacity based on the activation values of

the elements activated by the CS. To put it more precisely, the capacity is set to the sum of the activations of all elements of one particular CS *S*, where *S* is such that the sum of its elements' activations is higher than or equal to the sum of the elements' activations of each of the other CSs. Conceptually this amounts to the assumption that every single CS employed in the learning situation will always enter the attention buffer completely when presented alone (i.e., each CS can be fully attended). This is the "automatic" option. As an alternative, the buffer capacity can be set by hand to any real value above 0. The "fraction of common elements" parameter allows one to manipulate the overlap of elements representing different single CSs (for further details, see Harris, 2006). Possible values range from 0 to 1, where 0 indicates no overlap of CS elements (i.e., each CS activates its own elements only, and the elements of a certain single CS are activated only by this CS) and 1 indicates complete overlap of CS elements (i.e., each CS activates the same elements as every other single CS).

Since the way in which US and CS elements compete for entry into the attention buffer is not specified in Harris's present conception, using cutoff values in the learning equations (see Equation 2, p. 595, in Harris, 2006) is somewhat arbitrary. As a result, it is also in accord with the model to use a simple delta rule (assuming an activation of lambda for all US elements) for learning CS–US associations, thus neglecting buffer effects (for details, see Schultheis et al., 2008a). This parameter allows the user to specify whether the simulation should employ the simple delta rule or the more complex learning using cutoffs.

An Example

Here we demonstrate the usage of ALTSim with a simple summation experiment. All files needed, as well as further examples, can be downloaded from www.staff.uni-marburg.de/~lachnit/ALTSim/examples.

In a summation experiment, two components, A+ and B+, are reinforced separately. After successful training, A+, B+, and the compound AB? are tested. As we will see, the models differ in their predicted response strength for AB?, compared with the response strength for A+ or B+.

First, files containing the trial matrix and the stimulus code must be created. The trial matrix comprises the following lines:

A+	1	.8
B+	1	.8
AB?	0	0

In both reinforced component trials (A+ and B+), the US is present (lambda = 1), with a beta of .8. The lambda and beta of the AB? trial is arbitrarily set to 0; both values are not needed to compute the response to test trial AB?.

The stimulus code consists of two columns representing components A and B:

1	0
0	1
1	1

Component A is present in A+ trials (line 1) and AB? trials (line 3), but not in B+ trials (line 2). Component B is present in B+ and AB? trials, but not in A+ trials.

Both files have to be saved, for example, as summation_trialmatrix.txt and summation_stimuluscode.txt, respectively.

Now we can start the simulator using ALTSim.exe for the compiled version or by running ALTSim.m within MATLAB. The main GUI opens and we enter the file paths and check to see whether the files are read in correctly, by clicking the "View" button. Assuming that no stimulus sequences exist, we choose to generate sequences. More precisely, we would like to present 10 A+ and B+ trials each, in alternating order. Thus, we can use the default values for number of sequences and number of blocks and leave the randomize check box disabled. One training block then consists of one A+ and one B+ trial; therefore, we enter "1 2" in the training block field. Since we do not wish to save it, we leave the save sequence field empty. Now we click on the "1. initialize stimuli" button, and a message should tell us that the stimuli are initialized.

First, we would like to know what Pearce's configural model predicts, so we select it from the list box and click "2. choose theory." We are happy with using the default alpha values of 1 for each of the three configural units and do not wish to save anything, so we do not have to enter anything and can start the simulation by clicking the "3. simulate" button. A question dialog box will warn us that we did not enter the correct number of alphas, but we will not care because we are satisfied with the default value of 1. The graphic of Figure 1A appears, and we see that Pearce's configural model predicts that the response to AB? is the same as those to A+ and B+.

Now, let us see what REM predicts. We need to create a proportion file containing the following two lines:

.8
.8

The first value indicates that 80% of the elements of component A are replaced when A is presented together with B. The second value indicates that the same is true for B when A is present. After saving this matrix to a file (e.g., 2components.pro), we can enter the file path in the proportion matrix field of the REM GUI. Using the default value for alpha, we simulate the REM and see in Figure 1B that it predicts that the response to AB? is smaller than that to A+ or to B+ . By using the "view" button, we can alter the *r* values. An additional GUI will appear with the proportion matrix displayed in the lower panel. We change both values to .2, click "save," and start the simulation again. Examining the output activation graph (Figure 1C), we see that the REM now predicts that the response to AB? is larger than that to A+ or to B+. Therefore, the REM, but not Pearce's configural model, is able to explain diverging results of summation experiments. The Pearce model, however, is able to do this, too: A stimulus code including an additional component X, representing the context, produces a summation effect for

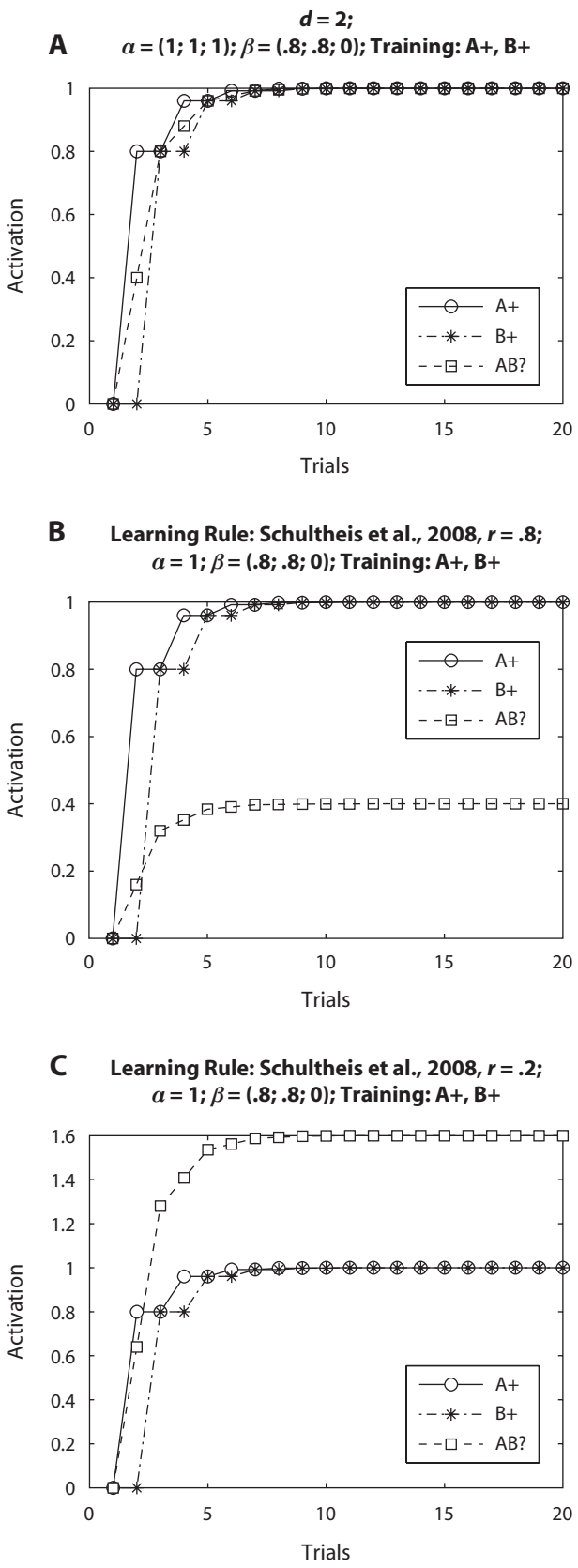


Figure 1. Simulated predictions for a summation experiment: (A) Pearce (1994), (B) REM with $r = .8$, and (C) REM with $r = .2$.

what will then be ABX? trials (for further analyses, see Glautier, Redhead, Thorwart, & Lachnit, 2008).

Conclusions

ALTSim provides an easy-to-use tool that facilitates the evaluation and comparison of different associative learning models. By means of ALTSim, we can gain a deeper understanding of the representations and processes involved in these models, we are able to design experiments, and we can explain existing empirical effects. Of course, more or other models could be included; the current version, however, already covers a wide spectrum, from elemental to configural models.

AUTHOR NOTE

This research was part of the doctoral thesis of the first author and was supported by the Deutsche Forschungsgemeinschaft (DFG) graduate program NeuroAct (DFG 885) and by a grant to the fourth author (DFG La 564/19-1). We thank Steven Glautier for lively and productive discussions concerning the implementation of the replaced elements model. Correspondence concerning this article should be addressed to A. Thorwart, Department of Psychology, Philipps-Universität Marburg, Gutenbergstraße 18, 35032 Marburg, Germany (e-mail: anna.thorwart@staff.uni-marburg.de).

REFERENCES

GLAUTIER, S. (2007). Simulation of associative learning with the replaced elements model. *Behavior Research Methods*, **39**, 993-1000.

GLAUTIER, S., REDHEAD, E., THORWART, A., & LACHNIT, H. (2008). *Reduced summation with common features in causal judgements*. Manuscript submitted for publication.

HARRIS, J. A. (2006). Elemental representations of stimuli in associative learning. *Psychological Review*, **113**, 584-605.

KINDER, A., & LACHNIT, H. (2003). Similarity and discrimination in human Pavlovian conditioning. *Psychophysiology*, **40**, 226-234.

LACHNIT, H., SCHULTHEIS, H., KÖNIG, S., ÜNGÖR, M., & MELCHERS, K. (2008). Comparing elemental and configural associative theories in human causal learning: A case for attention. *Journal of Experimental Psychology: Animal Behavior Processes*, **34**, 303-313.

MELCHERS, K. G., SHANKS, D. R., & LACHNIT, H. (2008). Stimulus coding in human associative learning: Flexible representations of parts and wholes. *Behavioural Processes*, **77**, 413-427.

PEARCE, J. M. (1987). A model for stimulus generalization in Pavlovian conditioning. *Psychological Review*, **94**, 61-73.

PEARCE, J. M. (1994). Similarity and discrimination: A selective review and a connectionist model. *Psychological Review*, **101**, 587-607.

REDHEAD, E. S., & PEARCE, J. M. (1995). Similarity and discrimination learning. *Quarterly Journal of Experimental Psychology*, **48B**, 46-66.

RESCORLA, R. A., & WAGNER, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. Black & W. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64-99). New York: Appleton-Century-Crofts.

SCHULTHEIS, H., THORWART, A., & LACHNIT, H. (2008a). HMS: A MATLAB simulator of the Harris model of associative learning. *Behavior Research Methods*, **40**, 442-449.

SCHULTHEIS, H., THORWART, A., & LACHNIT, H. (2008b). Rapid-REM: A MATLAB simulator of the replaced-elements model. *Behavior Research Methods*, **40**, 435-441.

WAGNER, A. R. (2003). Context-sensitive elemental theory. *Quarterly Journal of Experimental Psychology*, **56B**, 7-29.

WHITLOW, J. W., & WAGNER, A. R. (1972). Negative patterning in classical conditioning: Summation of response tendencies to isolable and configural components. *Psychonomic Science*, **27**, 299-301.

(Manuscript received June 7, 2008;
revision accepted for publication September 9, 2008.)

Running head: Addition of a common cue

Convergent results in eyeblink conditioning and contingency learning in humans: Addition of a common cue does not affect feature-negative discriminations

Anna Thorwart*

Philipps-Universität Marburg, Germany

Steven Glautier

University of Southampton, United Kingdom

Harald Lachnit

Philipps-Universität Marburg, Germany

* corresponding author

Addresses: Anna Thorwart & Harald Lachnit
Fachbereich Psychologie
Philipps-Universität Marburg
Gutenbergstraße 18
35032 Marburg, Germany
phone: +49 6421 2823715
fax: +49 6421 2826621
email: anna.thorwart@staff.uni-marburg.de
lachnit@staff.uni-marburg.de

Steven Glautier
School of Psychology
University of Southampton
Southampton
SO17 1BJ, UK
phone: (+44) 023-8059-2589
email: spg@soton.ac.uk

Abstract

We compared human discrimination learning in eyeblink conditioning and contingency learning. Participants had to learn two feature-negative discriminations: A+/ AB- as well as CD+/ CDE-. The Replaced Elements Model predicts that response differentiation will be facilitated by adding a common cue (C) to the reinforced and non-reinforced stimuli as done in the latter discrimination. The configural model of Pearce predicts impeded learning, whereas an extended version (including a discrimination parameter) predicts no difference. Convergent results were obtained in both paradigms: Adding a common cue did not affect response differentiation at all. These results add to earlier observations in human eyeblink conditioning which led to the suggestion that associative learning models have to take into account a discrimination parameter. Furthermore, they support the notion that learning in both paradigms is based on the same principles.

Convergent results in eyeblink conditioning and contingency learning in humans: Addition of a common cue does not affect feature-negative discriminations

The present experiments focus on the influence of a common cue on a feature-negative discrimination. In an ordinary feature-negative discrimination, a stimulus A is reinforced when presented alone (A+), but non-reinforced when presented together with a second stimulus B (AB-, with B called the feature-negative). Thus, the feature-negative discrimination equals the standard paradigm of conditioned inhibition, where B becomes inhibitory because its repeated presentation together with an excitatory stimulus is not followed by reinforcement. A similar discrimination is obtained by adding a common cue (C) to both kinds of trials, resulting in CD+ and CDE-, with C being irrelevant for the discrimination.

The Rescorla-Wagner Theory of associative learning (Rescorla & Wagner, 1972) easily can explain successful learning in both discriminations. However, it predicts faster discrimination learning in the “common cue” condition compared to an ordinary feature-negative discrimination. Pearce and Redhead (1993) took this as an example of the most counterintuitive predictions of the Rescorla-Wagner Theory. A+ versus AB- share one cue (A) while CD+ versus CDE- have two cues in common (C and D). Therefore, similarity should be higher in the latter discrimination and the intuitive assumption would be that it is harder to discriminate more similar stimuli. Nevertheless, the Rescorla-Wagner Theory predicts that increasing the similarity between two stimuli by adding a common cue to both *facilitates* discrimination of these stimuli. Pearce and Redhead examined this prediction in the feature-negative discriminations just described as well as in negative patterning discriminations. They found an opposite pattern of results: in an autoshaping paradigm pigeons learned slower in the “common cue” condition compared to the “standard feature-negative” condition. Chandra and Smith (1998) observed the same when they conditioned the proboscis extension reflex in honeybees. These results are crucial to the ongoing debate on elemental versus configural representations in associative learning models (for a review see Melchers, Shanks, & Lachnit, 2008) as they do not only contradict the

elemental model of Rescorla and Wagner (1972) and confirm the more intuitive prediction but are furthermore in line with Pearce's configural model (Pearce, 1987, 1994).

However, when Kinder and Lachnit (2003) investigated similarity and discrimination learning in human eyeblink conditioning with different negative patterning tasks, they found that the similarity between the stimulus compounds had no influence at all. Whereas Redhead and Pearce (1995) had found different learning rates for A/B/C+ versus ABC- and AB/AC/BC+ versus ABC- in pigeons, humans learned both discrimination problems at the same rate, despite the higher similarity between the binary compounds and ABC compared to the single stimuli and ABC.

One aim of the current paper therefore was to further investigate the influence of a common cue on discrimination learning in humans. In addition to investigating its influence in the feature-negative design for the first time in humans, we intended to compare results obtained in two different paradigms: eyeblink conditioning and contingency learning. Despite the previous replications of experiments of one paradigm in the other, this is the first time that both were tested in parallel, adapting the procedures as far as possible and reasonable, and it should allow us to draw conclusions on the generality of our findings and of the rules that govern discrimination learning. For example, Lachnit, Schultheis, König, Üngör and Melchers (2008) used the same discrimination design in a contingency learning experiment as Kinder and Lachnit (2003) in their eyeblink study, but observed that discrimination of A+/B+/C+ versus ABC- was learned slower than AB+/AC+/BC+ versus ABC-. Because of changes to the stimulus material and procedures it is, however, not clear whether the divergent results within human learning had been due to fundamental differences between eyeblink conditioning and contingency learning or due to the other changes that had been made.

The second aim of the paper concerns the theoretical models. In the last decade, extensions and derivatives of both the Rescorla-Wagner Theory and Pearce's configural model were proposed. In order to explain their results, Kinder and Lachnit (2003) suggested an extension of Pearce's configural model by introducing a

discrimination parameter *d* that reflects the general discriminability of the stimulus material used in the specific task. This extension of Pearce's configural model was later also used by Pearce, Esber, George, and Haselgrove, (2008). On the other side, Wagner and colleagues developed a derivative of the Rescorla-Wagner Theory, the Replaced Elements Model (Brandon, Vogel, & Wagner, 2000; Wagner, 2003). In contrast to the Rescorla-Wagner Theory, stimulus representation is more complex here and may change depending on the context of presentation. Representational elements activated by stimulus A differ when A is presented alone compared to when A is presented together with a second stimulus, B. As the feature-negative designs so far have been successfully used to evaluate the precursor models, we decided to use them in the evaluation of the up-to-date models, too.

Glautier (2008b) already suggested that there are differential theoretical predictions from the Replaced Elements Model and Pearce's configural model. Figure 1 shows simulation results for the ordinary (A+ AB-) and the common cue (CD+ CDE-) discrimination. Simulations were computed using ALTSim 3.5 (Thorwart, Schultheis, König & Lachnit, 2009; see also Glautier, 2007). In the upper panels, predictions of the extended configural model are displayed. Pearce's configural model and its extension both assume that the whole stimulus compound presented in a learning trial (e.g. CDE) is represented by one configural unit and learning involves the association of this configural unit with the outcome unit. However, discrimination learning is impeded by generalisation between the to-be-discriminated stimulus compounds and generalisation increases with the similarity between the two compounds. Since both configural models define similarity in terms of the number of common cues, the similarity between CD+ and CDE- is higher than that between A+ and AB- as mentioned before. This results in more generalisation of excitation from CD to CDE than from A to AB (and vice-versa for generalisation of inhibition). Therefore, Pearce's configural model, that corresponds to the extended configural model with $d = 2$ (upper left panel of Fig. 1), predicts that adding a common cue should result in impeded learning. The right panel shows the predictions of the extended configural model with *d* increased to 20. An increasing *d* reflects increasing discriminability or decreasing generalisation

between the stimuli. With decreased generalisation, the effects just described diminish, in the limit to zero, so that both feature-negative discriminations are learned at the same rate.

Fig. 1 about here

The lower panel shows the predictions of the Replaced Elements Model. Like the Rescorla-Wagner Theory, the Replaced Elements Model assumes that a stimulus is decomposed into its parts and that these parts are represented and gain associative strength separately. In the Replaced Elements Model, however, contrary to the Rescorla-Wagner Theory a stimulus A is represented by different representational elements depending on the context of its presentation (e.g., whether a stimulus B is present or not). Wagner (2003) conceptualized the proportion of context-dependent elements in the total representation of a stimulus as a free parameter, *r*, that reflects properties like the perceptual interaction between, in our example, A and B. With no context-dependent elements ($r = 0$), the predictions of the Replaced Elements Model equal those of the Rescorla-Wagner Theory. With $r = 1$, the representation of A comprises completely different elements within different contexts. In spite of this variability, the Replaced Elements Model is not able to adapt its predictions in the same way as the extended configural model, although its parameters influence the amount of generalisation. The lower panels of Figure 1 show simulations for the two extreme values of *r*. The predictions in both cases, however, are quite similar: learning about CD+ versus CDE- is faster than learning about A+ versus AB-. This is true for the Replaced Elements Model for the same two reasons as for the Rescorla-Wagner Theory. First, CD will gain more excitatory associative strength than A in the first reinforced trial because - assuming same learning parameters for all stimuli - the initial growth in the strength of the CD compound ($V_{CD} = V_C + V_D$) will be twice as strong as that of the single stimulus A (V_A). Second and consequently, the excitatory strength of CDE after the first reinforced trial is greater than that of AB, and this will support faster

inhibitory learning about E than about B in the non-reinforced trials. The net result is that the predicted associative strengths of the CD and CDE compounds will diverge more rapidly than those of A and AB. In the Replaced Elements Model, it will also always be the case that the more stimuli are present, the more representational elements will be activated. And the effect of more activation in the "common cue" discrimination will always enhance learning.

To test these three different predictions (impeded versus equal versus facilitated learning), we compared the course of response differentiation in the two feature-negative discriminations described above using two well established paradigms of human learning. In Experiment 1, we used eyeblink conditioning (as Kinder and Lachnit, 2003). In Experiment 2, the task and stimulus material of Experiment 1 was used in a contingency learning experiment.

Experiment 1

Method

Participants. Twenty-seven students of the University of Marburg took part for course credit or payment. Two of them had to be excluded due to corrupt data files. Among the remaining 25 (23 females) ages ranged from 19 to 45 (M=21.8). None of them showed habituation of the unconditioned response (UR) in course of the experiment (more than 10 % of the reinforced trials without UR).

Stimuli. Conditioned stimuli (CSs) were nine coloured circles that could each appear in one of nine positions within a black rectangle on the lower half of the screen. Colour and position of the circles were assigned to stimuli A – I randomly for each participant. The remainder of the screen was coloured light blue to match the experimental setting of the contingency learning study in Experiment 2.

The unconditioned stimulus (US) was a corneal air puff. The air pressure was adjusted individually for each participant before the experiment so as to elicit an eyeblink but not to be painful.

Apparatus. The participants were seated 70 cm in front of a computer display showing the CSs. Conditioning of the eyeblink response was conducted at the left eye of the participant. The participant wore a headband with the air jet and the apparatus measuring the response attached to it. The latter was a piano wire tapped to the participant's eyelid transmitting movement to a photoelectric transducer.

Procedure. The participants were told that they would see stimuli on the screen and experience air puffs to their left eye. After adjusting the air puff, they were instructed to concentrate only upon the patterns on the display and to refrain from thinking of their eyeblink responses. No further explanations were given.

The experiment consisted of 144 trials, 24 trials each of A+, AB-, CD+, CDE-, F-, GHI+. F- and GHI+ were included as fillers so that the number of stimuli presented in compound was not predictive of the US. The sequence was pseudo-randomised for each participant with a maximum of three trials of same contingency following each other, structured in blocks of two trials per type.

The CS lasted for 900 ms; its offset was immediately followed by the US. The ITI varied randomly between 5 to 7 s with an average of 6 s.

Dependent variables. Response in the interval between 250 and 900 ms after CS onset were counted as conditioned responses (CRs), thus excluding alpha response as well as USs. Splitting the experiment into six blocks, proportions of trials with CR were calculated for each CS separately and a square root transformation of the proportions was performed to approximate normality of the positively skewed distributions.

Result and Discussions

For all statistical analyses, α was set to .05. The Huynh-Feldt correction of degrees of freedoms was used where appropriate.

Fig. 2 about here

Figure 2 shows the response frequencies for both feature-negative tasks. It is evident that both discriminations were learned successfully without any difference between the ordinary and the "common cue" feature-negative task. This was confirmed by a 2*2*6 ANOVA including the within-subject factors task (ordinary vs. common cue), contingency (reinforced vs. non-reinforced), and block (1-6). It revealed a significant main effect of contingency ($F(1, 24)=15.23$, $p < .002$, $\eta_p^2=.388$) and of block ($F(5, 120)=6.21$, $p < .001$, $\eta_p^2=.206$), as well as a significant Contingency*Block interaction ($F(5, 120)=5.32$, $p < .002$, $\eta_p^2=.18$). No effect including task was significant (all Fs < 1, $\eta_p^2<.035$).

Was our failure to find differences in response differentiation between the two tasks due to a lack of power? Most relevant in respect to the predictions is the Contingency*Task*Block interaction. The average observed correlation between all measurements was .32, Huynh-Feldt's nonsphericity correction ε was 1. The resulting power to find a medium effect of $f = .25$ was 99.9 %, and the power to find a small effect of $f = .1$ still was 59 %.

Experiment 2

In Experiment 1 participants learned feature-negative discriminations in eyeblink conditioning, but adding a cue made the task neither more difficult nor easier. In Experiment 2, we employed a contingency learning paradigm and similar procedures to assess the generality of our conditioning results. Furthermore, with contingency learning a more detailed inspection of early training is possible as in contrast to eyeblink conditioning blocking of several trials together is not necessary. The experimental setting was the same as in the eyeblink conditioning study besides three changes. First, some elements were added to the background in order to deliver the impression of a flight control centre, matching the cover story. Second, to investigate the underlying learning processes, a summation test of inhibition was included for the "feature-negatives" B and E, and necessarily, a reinforced stimulus F+ was presented during training. Third, we increased the number of participants to further enhance the test power for the relevant triple interaction.

Method

Participants. Fifty-two participants were recruited from the University of Southampton Highfield Campus. The ages ranged from 18 to 46 years (M=22.2) and there were 30 females. Participants were paid £4 or received course credit.

Apparatus and Stimuli. A computer display delivered the impression of a flight control centre (see Fig. 3). The CSs were the same as in Experiment 1. The outcome was the pollution level produced by a plane which flew across the computer screen. The pollution level was displayed continuously as a number above the coloured circles and updated approximately every 0.25 s. It varied randomly around an average of 20 (range 15-25) except when the positive outcome (i.e. the US) was scheduled. The positive outcome consisted of an increase in pollution to an average of 40 (range 35-45).

Fig. 3 about here

Procedure. The task for the participants was to predict, on the basis of the coloured circles, when the pollution level would rise (see Appendix). Each trial began with a plane flying across the screen and a CS coming on. The CSs lasted for 3 s. If a positive outcome was scheduled to occur, it happened 1 s after CS onset and lasted 2 s. There was a 4 s ITI.

During the experiment participants experienced 12 training trials each of A+, AB-, CD+, CDE-, F+, H-, I-, GHI+ (the latter three serving as filler cues). They were arranged in blocks and after each block there was a test stage. The first four blocks had one trial of each type to facilitate identification of early learning. The final four blocks had two trials of each type. During tests, training was interrupted and participants indicated what pollution level they would expect for the following cues or compounds: A, AB, CD, CDE, F, BF, and EF. Due to technical problem, eleven participants did not receive test trials with EF. All participants had time as long as needed to respond. The

order of test trials was determined randomly. Similarly, the order in each training block was randomised.

Result and Discussions

Fig. 4 about here

The left panel of Figure 4 shows the ratings for both feature-negative discriminations. Whereas both discriminations were learned successfully in the course of the experiment, no effect of the common cue was apparent. This was confirmed by a 2*2*8 ANOVA including the within-subject factors task (ordinary vs. common cue), contingency (reinforced vs. non-reinforced), and block (1-8). It revealed a significant main effect of contingency ($F(1, 51)=8.11$, $p < .007$, $\eta_p^2=.137$) and of block ($F(7, 357)=2.45$, $p < .037$, $\eta_p^2=.0.46$), as well as a significant Contingency*Block interaction ($F(7, 357)=6.29$, $p < .001$, $\eta_p^2=.11$). Again, none of the effects including task (Fs< 1.21, $p>.296$, $\eta_p^2<.03$) was significant. Due to the increased number of participants (52 in Exp. 2 compared to 25 in Exp. 1) the test power for the relevant triple interaction was now 88 % for detecting a small effect ($f = .1$, average $r = .11$, $\varepsilon = .931$).

The results of the summation test of inhibition are displayed in the right panel of Figure 4. The rating of the reinforced single stimulus F+ increased during the experiment whereas the rating of the test compounds BF and EF stayed about the same level, indicating that B and E became more inhibitory across training. A comparison of the ratings of BF and EF in a 2*8 ANOVA (BF vs. EF * block 1-8) showed no significant effects. All Fs but that for the Task*Block interaction ($F < 1.3$, $p > .276$, $\eta_p^2 < .03$) were less than 1.[1]

The results of this experiment clearly replicated those of the eyeblink conditioning study: Both feature-negative discriminations were learned at the same rate. In contingency learning this was evident in the feature-negative discriminations themselves as well as in the summation test of inhibition for both “feature-negatives”.

Inhibitory learning of the feature-negative in the “common cue” task (E) was not faster than inhibitory learning in the ordinary task (B). This, together with the sufficient test power for small effects, establishes a rather firm basis for the conclusion that adding a common cue to a feature-negative discrimination does not enhance the amount of response discrimination in human contingency learning.

General Discussion

We compared human discrimination learning in two feature-negative tasks employing eyeblink conditioning and a contingency learning paradigm in humans. The Rescorla-Wagner Theory (Rescorla & Wagner, 1972) and the Replaced Elements Model (Wagner, 2003) predict that learning of a feature-negative discrimination should be enhanced by increasing similarity, whereas Pearce’s configural model (Pearce, 1994) predicts a disruption. Therefore, investigating the influence of similarity on learning is crucial for the debate on elemental versus configural stimulus representations within associative learning. In both experiments, however, we found neither an enhancement nor a disruption. Therefore, the predictions of none of the three theories mentioned were in accordance with the data. Instead, the extension of Pearce’s configural model proposed by Kinder and Lachnit (2003) successfully predicts the observed data, if a high discrimination parameter ($d > 20$), i.e. little generalisation, is assumed. Kinder and Lachnit initially had suggested this extension in a post hoc fashion in order to deal with their observations in negative patterning tasks. Although they had manipulated the similarity between reinforced and non-reinforced cues the difficulty of those discriminations was not influenced at all. The present experiments replicated this observation and therefore provide an empirical proof of the suggested extended configural model. One might argue that this support rests upon the acceptance of the null hypothesis. The predicted null result, however, was replicated in two experiments with considerably high power. These results therefore might well indicate that humans compared to other animals generalise very little between stimuli in feature-negative and negative patterning tasks, whereas pigeons and bees learn more slowly with increasing similarity (Chandra & Smith, 1998; Deisig, Lachnit, Sandoz, Lober, & Giurfa,

2003; Pearce & Redhead, 1993; Redhead & Pearce, 1995).

Importantly, this seems to be a rather general principle in human learning as in the present experiments eyeblink conditioning and contingency learning yielded identical patterns of results. At a first glance, this seems to be in contrast to the findings of Lachnit et al. (2008) with a negative patterning task reported in the introduction. In a contingency learning experiment they used the same discrimination design as Kinder and Lachnit (2003) had used in eyeblink conditioning. Instead of coloured squares, however, Lachnit et al. used pictures of fruit and vegetables to meet the requirements of their cover story. With a food allergy task they observed that the discrimination of A+/B+/C+ versus ABC- was learned slower than the discrimination of AB+/AC+/BC+ versus ABC- , a finding being in line with elemental models. Lachnit (1988) already had shown that the kind of stimulus material clearly influences the kind of stimulus processing, and Glautier (2002) demonstrated that even the spatial arrangement of the stimuli on the screen can have an effect. Hence, the different response patterns might have been due to differences in stimuli. On the other hand, Kinder and Lachnit used three clearly discriminable lines as stimuli in their first experiment and a large number of red, yellow, and green rectangles in their second and third experiment, but nevertheless observed the missing influence of similarity in all three experiments. To complicate things even more, in another causal learning experiment, Lachnit et al. (2009) used the stimuli of Kinder and Lachnit's Experiments 2 and 3. This way, they replicated the results of the previous causal learning experiments with pictures of fruits, but not those of Kinder and Lachnit with coloured rectangles. Therefore, differences in stimuli are not a plausible explanation for the deviating observations made so far.

Yet, there is another aspect in which the former causal learning and eyelid conditioning experiments clearly differed but the current ones matched. In each trial of the causal learning study of Lachnit et al. (2008) participants had to make explicit judgments about whether or not they expected that certain food leads to an allergic symptom, and only after this rating the correct outcome was presented. In eyeblink conditioning as well as in the present contingency learning experiment, on the other

hand, no explicit response was required during training but the CS and the outcome (or US) were presented in a constant stream (in the contingency learning task interrupted six times by the test stages). With respect to causal learning and recency and primacy effects, other studies have shown an effect of judgement frequency (Matute, Vegas, & De Marez, 2002; Glautier, 2008a) and this could also influence discrimination learning.

In sum, when in the present experiments we held constant the discrimination task, the stimuli, and - as far as reasonable - the experimental procedures, we obtained identical patterns of results in human eyeblink conditioning and in contingency learning. This supports the idea that both learning tasks are based upon the same principles. Furthermore, the present results replicated the findings of Kinder and Lachnit (2003) in a different discrimination task and therefore support the extended version of Pearce's configural model.

References

Brandon, S. E., Vogel, E. H. & Wagner, A. R. (2000). A componential view of configural cues in generalization and discrimination in Pavlovian conditioning. *Behavioural Brain Research, 110*, 67-72.

Chandra, S., & Smith, B.H. (1998). An analysis of synthetic processing of odor mixtures in the honeybee (apis mellefera). *Journal of Experimental Biology*, *201*, 3113-3121.

Deisig, N., Lachnit, H., Sandoz, J.-C., Lober, K. & Giurfa, M. (2003). A modified version of the unique-cue theory accounts for olfactory compound processing in honeybees. *Learning and Memory, 10*, 199-208.

Glautier, S. (2002). Spatial separation of target and competitor cues enhances blocking of human causality judgements. *Quarterly Journal of Experimental Psychology, 55B*, 121-135.

Glautier, S. (2007). Simulation of associative learning with the replaced elements model. *Behavior Research Methods, 39*, 993-1000.

Glautier, S. (2008a). Recency and primacy in causal judgments: effects of probe question and context switch on latent inhibition and extinction. *Memory and Cognition, 36*, 1087-1093.

Glautier, S. (2008b). The elemental-configural distinction: A problem of two dimensions Commentary on Melchers, Shanks, and Lachnit (2007). *Behavioural Processes, 77*, 431-433.

Kinder, A. & Lachnit, A. (2003). Similarity and discrimination in human Pavlovian conditioning. *Psychophysiology, 40*, 226-234.

Lachnit, H. (1988). Convergent validation of information processing constructs with Pavlovian methodology. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 143-153.

Lachnit, H.; Schultheis, H.; König, S.; Üngör, M. & Melchers, K. (2008). Comparing elemental and configural associative theories in human causal learning: A case for attention. *Journal of Experimental Psychology: Animal Behavior Processes*, *34*, 303-313

Lachnit, H., Thorwart, A., Schultheis, H., Uengoer, M., Lotz, A. & König, S. (2009). *Indicators of early and late processing reveal the importance of within-trial-time for theories of associative learning.* Manuscript submitted for publication.

Matute, H., Vegas, S. & DeMarez, P.-J. (2002). Flexible use of recent information in causal and predictive judgments. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 28*, 714-725.

Melchers, K. G., Shanks, D. R. & Lachnit, H. (2008). Stimulus coding in human associative learning: Flexible representations of parts and wholes. *Behavioural Processes, 77*, 413-427.

Pearce,J.M. (1987). A model of stimulus generalization for Pavlovian conditioning. *Psychological Review*, *94*, 61-73.

Pearce, J. M. (1994). Similarity and discrimination: A selective review and a connectionist model. *Psychological Review, 101,* 587-607.

Pearce, J. M., Esber, G. R., George, D. N., & Haselgrove, M. (2008). The nature of discrimination learning in pigeons. *Learning & Behavior*, *36*, 188-199.

Pearce, J.M. & Redhead, E. (1993). The influence of an Irrelevant Stimulus on Two Discriminations. Journal of Experimental Psychology. *Animal Behavior Processes*, *19*, 180-190.

Redhead, E. S. & Pearce, J. M. (1995). Similarity and discrimination learning. *Quarterly Journal of Experimental Psychology*, *48B*, 46-66.

Rescorla, R., & Wagner, A. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In A. Black & W. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64-99). New York: Appleton-Century-Crofts.

Thorwart, A., Schultheis, H., König, S., & Lachnit, H. (2009). ALTSim: A MATLAB simulator for current associative learning theories. *Behavior Research Methods, 41*, 29-34.

Wagner, A. R. (2003). Context-sensitive elemental theory. *Quarterly Journal of Experimental Psychology, 56B*, 7–29.

Author Note

Anna Thorwart and Harald Lachnit, Department of Psychology, Philipps-Universität, Marburg, Germany. Steven Glautier, School of Psychology, Southampton University, UK.

Anna Thorwart is supported by the DFG (Deutsche Forschungsgemeinschaft) graduate program NeuroAct (DFG 885) and Harald Lachnit by grant La 564/19-1.

Correspondence concerning this article should be addressed to Anna Thorwart, Department of Psychology, Philipps-Universität Marburg, Gutenbergstraße 18, 35032 Marburg, Germany (email: anna.thorwart@staff.uni-marburg.de).

Footnotes

[1] If one computes statistical test for the single trials, despite the non-significant overall ANOVA, the t-test for the difference between EF and BF in the third trial revealed the following statistics: $t(40)=2.886$, $p<.01$, $d=.45$. This is not significant if one applies a Bonferroni Correction for the eight t-tests. Furthermore, neither A+ and CD+ nor AB- and CDE- differed during the whole experiment. It is not clear why E should have larger inhibitory power towards F than towards CD which were rated equally. E.g., there could not be any ceiling effect for the later as the ratings were far from asymptotic.

Figure Captions

Figure 1. Simulations of the feature-negative discriminations. The upper two panels display predictions of the extended configural model of Pearce: in the left panel, the generalisation is computed as in the original model of Pearce (1994, discrimination parameter *d* =2); in the right panel, generalisation is reduced by enhanced discrimination (*d* = 20). The lower panels display predictions of the Replaced Elements Model with different proportion of context-dependent elements (*r*): the left panel with *r* = 0 equals the Rescorla-Wagner Theory, in the right panel all elements are context-dependent (*r* = 1).

Figure 2. Mean responding (± 1S.E.) in Experiment 1 to the ordinary feature-negative task (A+ vs. AB-) and the "common cue" task (CD+ vs. CDE-)

Figure 3. Example of screen in Experiment 2. The coloured circles served as CS, the number displayed above represents the air pollution level.

Figure 4. Mean ratings (± 1S.E.) in Experiment 2.The first four tests were after blocks of one training trial per type, the remaining four after blocks of two. The left panel includes the rating of both feature-negative discriminations, the right the summation test of inhibition for both feature-negatives (B and E).

fig1

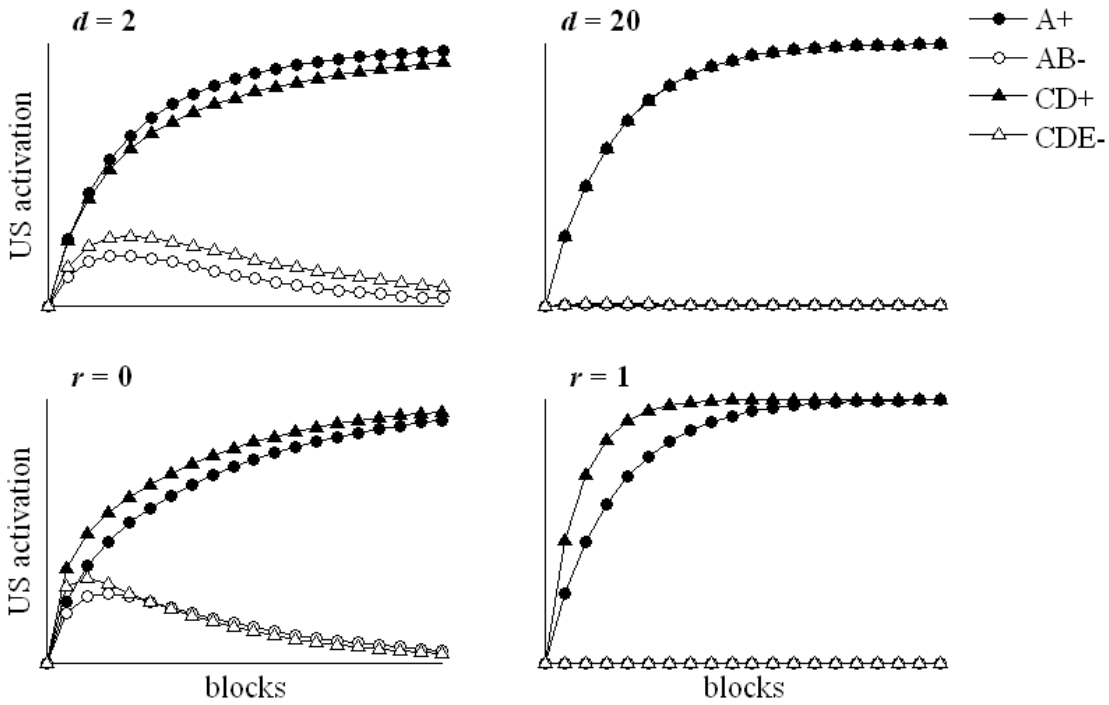

d = 2
d = 20
r = 0
r = 1
US activation
US activation
blocks
blocks
A+
AB-
CD+
CDE-

fig2

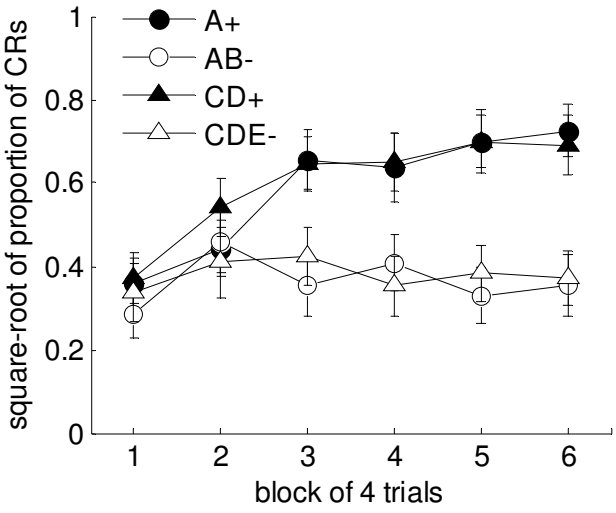

A+
AB-
CD+
CDE-
square-root of proportion of CRs
block of 4 trials
0
0.2
0.4
0.6
0.8
1
1
2
3
4
5
6

fig3

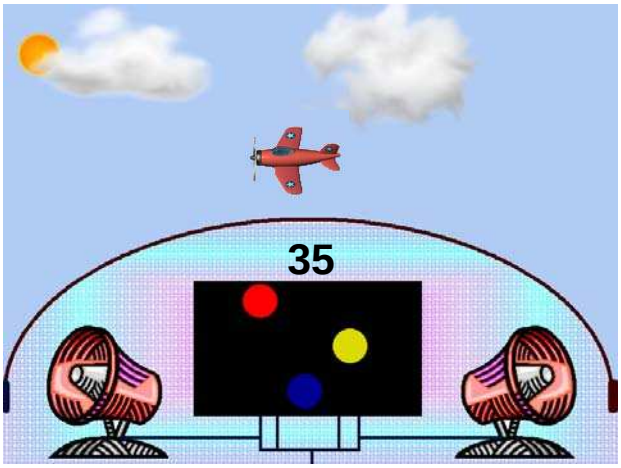
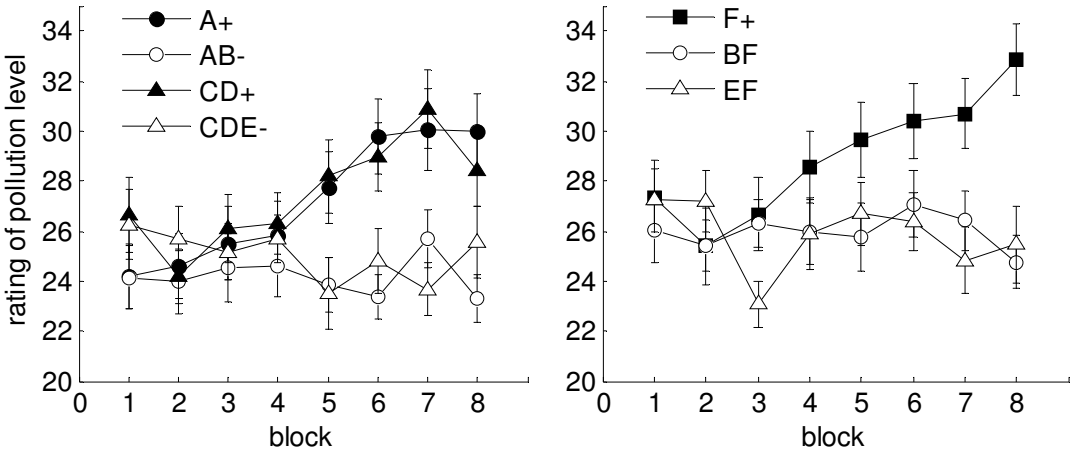


fig4


A+
AB-
CD+
CDE-
F+
BF
EF
rating of pollution level
block
block
20
22
24
26
28
30
32
34
0
1
2
3
4
5
6
7
8

Appendix

Instructions for participants in Experiment 2.

During this experiment you will see a plane "flying" across the computer screen. Below the plane is a display of its instrument panel. The instrument panel gives a continuous display of the pollution level that is produced by the plane's engine. High pollution levels are critical because they breach health and safety regulations. The instrument panel also provides visual signals that might indicate that changes in pollution levels are about to occur. Your job is to learn the extent to which different signals can be used to predict pollution levels. What you have learned about the different signals will be tested at various points during the experiment. During a test trial the display of the plane will stop and you will be presented with instructions for making your rating. When you have read the instructions press a key and a signal will be presented. Once the signal has been presented you will be asked to indicate the level of pollution you would expect to follow that signal. Some test signals might be ones that you haven't experienced before. Use your experience to estimate what you think the pollution level would be. Once you have made your rating the experiment will continue when you press the RETURN key. When you have read and understood these instructions press a key to continue.

Learning & Behavior
2009, 37 (1), 107-115
doi:10.3758/LB.37.1.107

Symmetrical generalization decrements: Configural stimulus processing in human contingency learning

Anna Thorwart and Harald Lachnit
Philipps-Universität Marburg, Marburg, Germany

Models of associative learning differ in their predictions concerning the symmetry of generalization decrements. Whereas Pearce's (1994) configural model predicts the same response decrement after adding elements to and after removing elements from a previously trained stimulus, elemental models, such as the replaced elements model and Harris's (2006) model, anticipate more of a decrement for removing than for adding elements. In three contingency learning experiments, we manipulated the motion and the spatial arrangement of colored dots in order to induce configural or elemental processing by perceptual grouping. The results reliably showed symmetrical decrements for the *added* and *removed* groups. The manipulations of the stimuli had no effect on stimulus processing. This is in line with Pearce's configural model, but it is at variance with the elemental models and previous studies.

Stimulus generalization occurs when a response to a new and neutral stimulus resembles a conditioned response (CR) to a previous conditioned stimulus (CS). After eating a poisonous apple, we presumably will experience nausea not only if we taste the same kind of apple again, but also if we taste a similar kind of apple. Knowledge about the CS—the poisonous apple—generalizes to a new stimulus—the other apple. Generalization between stimuli is a vital ability because it allows for probably appropriate responding to new stimuli. Responding to the new stimulus, however, will normally be weaker than responding to the trained stimulus. This difference in response strength is referred to as *generalization decrement*.

Models of associative learning explain generalized responding to an apparently new stimulus by assuming that this new stimulus will share components[1] with the previously conditioned CS. This is true for so-called *elemental* and *configural* models. Although previous research has often tried to ascertain whether processing is configural or elemental, the issue has now shifted toward identifying the conditions under which processing is best described by configural theories and those under which it is best described by elemental theories (Melchers, Shanks, & Lachnit, 2008, with commentaries; Shanks, Lachnit, & Melchers, 2008). Distinctions among associative learning models are based on differences in how these models conceptualize the coding, processing, and representation of stimuli. In configural models like Pearce's (1987, 1994), a stimulus compound is represented and associated as a whole. Generalization is based on the coactivation of representations of similar compounds, with the amount of coactivation depending on the proportion of common components. The central characteristic of elemental models, on the other hand, is the assumption that individual components of a stimulus compound are learned about. Stimuli with common components will therefore activate the same associations. Modern elemental models, such as the replaced elements model (REM; Wagner, 2003) and Harris's (2006) model, assume flexible nonlinearity, which means that a compound need not activate exactly the same associations as its components.[2]

In the following studies, we compared the application of configural models and of elemental models to two different generalization tasks. In the first task, the new stimulus (test stimulus) was created by removing some components of the just-trained CS (e.g., training AB+ and testing A). This course of action is equivalent to the design of *overshadowing* experiments. In the second task, after training the CS (e.g., A+), some new components are added (e.g., AB). If responding to the test stimulus AB is weaker than responding to the initially trained A, this can be described as *external inhibition*. Both overshadowing and external inhibition are well-known phenomena in the study of associative learning (Pavlov, 1927).

We simulated the predictions of the different models by using several MATLAB-based programs (Schultheis, Thorwart, & Lachnit, 2008a, 2008b; Thorwart, Schultheis, König, & Lachnit, in press). Pearce's (1994) configural model predicts symmetrical decrements; that is, the amount of decrement produced by testing AB after training with A will be the same as that produced by testing A after training with AB. The decrement is based on the coactivation of similar configural units, and just as similarity is symmetric, so too are generalization decrements. On the

A. Thorwart, anna.thorwart@staff.uni-marburg.de

other hand, both elemental theories predict asymmetrical generalization decrements: Removal of components always results in a larger decrement than does adding them. Interestingly, the flexibility of the elemental models under consideration does not allow us to predict symmetrical decrements. In sum, comparing generalization decrements after adding and after removing components theoretically clears the way for contrasting the predictions of Pearce's (1994) configural model with those of Wagner's (2003) and Harris's (2006) elemental models.

Up to now, studies that have investigated generalization decrement explicitly have routinely observed asymmetrical generalization decrements (Brandon, Vogel, & Wagner, 2000; Glautier, 2004; González, Quinn, & Fanselow, 2003; Wheeler, Amundson, & Miller, 2006). Brandon et al. trained rabbits in an eyelid-conditioning study with A+, AB+, or ABC+. Using cross-modal stimuli (light, tone, vibration) and testing with A, AB, and ABC, Brandon et al. observed larger generalization decrements after removing than after adding a stimulus component. González et al. conditioned rats across 2 consecutive days. A single unsignaled footshock was presented in the presence of specific contextual cues (light, tone, odor). Using freezing as a measure of conditioning, removal—but not addition—of a stimulus component of the training context produced a significant generalization decrement. Glautier (2004) and Wheeler et al. used causal learning tasks to investigate generalization decrements. In Glautier's (2004) experiments, participants had to rate the amount of air pollution produced by airplanes. Stimuli created by adding novel features (such as markings or guns) to the previously trained airplanes were rated at the same level as the original training stimuli, but stimuli created by removing features led to reduced ratings. Wheeler et al., in a food-poisoning scenario, observed that adding a labeled picture of food to a previously learned picture of food can produce a generalization decrement, but removing produces a more robust decrement. Although these studies differed in the size of decrements produced by adding a stimulus component, they nevertheless converged in the observation that removing stimulus components produces larger decrements. Therefore, all of these authors concluded that elemental models were better than configural ones in describing stimulus processing during stimulus generalization.

In spite of these findings, there may yet be circumstances in which configural processing operates. For other associative learning tasks, such as summation or blocking, it has been shown that results consistent with both configural and elemental accounts can be obtained (Melchers et al., 2008). In an attempt to investigate whether manipulation of the stimuli is able to produce symmetrical generalization decrements, we addressed the question, "Under which conditions is processing best described by configural theories, and under which conditions is it best described by elemental theories?" One successful attempt in this line concerned the spatial arrangement of the stimulus components in blocking experiments (Glautier, 2002). Blocking was reduced when the same components were presented in close proximity, whereas spatial separation of the components enhanced blocking.

A common explanation relates perceptual grouping and gestalt principles to associative learning. The term *perceptual grouping* describes the notion that various elements in a complex display may be perceived as "going together" (Palmer, 1999). It seems reasonable that objects seen as belonging together perceptually may be processed configurally. Thus, the manipulation of perceptual grouping of the components may affect processing in associative learning. According to the gestalt principle of proximity, objects that are closer together will be seen as belonging together (Palmer, 1999). This would explain the results of Glautier (2002). If this is the case, other gestalt principles too should have a comparable impact on stimulus processing and, by this, on the symmetry of generalization decrements.

Common fate is another powerful principle, according to which objects moving with the same speed in the same direction become perceptually grouped and are seen as belonging together (Palmer, 1999). Therefore, we used this principle in an attempt to induce configural processing in a stimulus-generalization study.

EXPERIMENT 1

In the present experiment, small dots differing in color were seen moving across a black screen for 2 sec. The kind of motion varied as a dichotomous between-participants factor: In the *separated motion* conditions, the direction of motion was randomly chosen for each dot and trial, with some restrictions; in the *grouped motion* conditions, all dots were moving in the same direction, chosen randomly for each trial. To compare the decrements, the second between-participants factor *task* was included: *Added* groups were trained with two stimulus configurations (one reinforced and the other nonreinforced), consisting of two different-colored dots each. Thus, color was the relevant feature, whereas motion was irrelevant for the outcome. For *removed* groups, the two training stimuli (one reinforced and the other nonreinforced) consisted of five different-colored dots each. In the test stage after training, the five- and the two-dot compounds were presented in both tasks. This two-factor design resulted in four experimental groups: *separated–added*, *separated–removed*, *grouped–added*, and *grouped–removed* (see Table 1). The qualitative patterns of the predictions already have been described above: If perceptual grouping can induce configural processing, generalization decrements should be asymmetrical in the separated condition, but symmetrical in the grouped condition.

Method

Participants. We recruited 140 participants from the Philipps-Universität Marburg campus by posted advertisement and by face-to-face contact. They received course credits or were paid with sweets. We excluded 44 participants from further analysis because they did not solve the discrimination at the end of the training. To be included, the rating of the reinforced stimulus had to be higher than 4, and that of the nonreinforced stimulus had to be lower than 4, in the last trial of training; these numbers refer to scores on a 7-point scale (see below). The remaining 96 participants were from 18 to 48 years of age (M = 23.69). Sixty-nine women took part. We tested 20, 25, 30, and 21 participants in the grouped–removed, separated–removed, grouped–added, and separated–added groups, respectively.

Table 1
Design of Experiments 1, 2, and 3

	Group			Test Stage	
Experiment	Stimuli	Task	Training Stage	Trained Stimuli	Test Stimuli
1	separated grouped	added	AB+, QR−	AB, QR	ABCDE, QRSTU
	separated grouped	removed	ABCDE+, QRSTU−	ABCDE, QRSTU	AB, BC, CD, DE, AE, QR, RS, ST, TU, QU
2	—	added	AB+, QR−	AB, QR	ABC, QRS
		removed	ABC+, QRS−	ABC, QRS	AB, BC, AC, AB, AC, QR, RS, QS, QR, RS
3	variable fixed	(*within-subjects*)	AB+, CD−, EFG+, HIJ−	AB, CD, EFG, HIJ	ABX, CDY, EF, IJ

Note—Each letter refers to a dot with a specific color. A "+" indicates that the temperature was above the critical limit (reinforced trials), and a "−" indicates that the temperature was below the critical limit (nonreinforced trials). The numbers of trials are not specified (see text for details).

Apparatus and Stimuli. The experiment took place either in our psychology labs or in the reading room of the university's student services house. In the labs, the stimuli were presented on 15-in. color monitors with 1,024 × 768 resolution and a 75-Hz refresh rate. In the student services house, an IBM ThinkPad was used, placed in a paperboard cabin. The monitor was 15 in., with 1,024 × 768 resolution and a 60-Hz refresh rate. The participants wore sound-attenuating earmuffs.

Different-colored dots were used as the CS. The size of the dots was 16 pixels. Ten different colors were chosen (RGB: 64, 204, 255; 53, 169, 55; 234, 50, 24; 147, 36, 207; 238, 29, 90; 187, 212, 65; 231, 95, 217; 69, 49, 181; 255, 157, 0; 253, 228, 111). The colors were assigned randomly for each participant but, overall, the separated and grouped conditions were matched in the colors used.

The dots moved across a central 600 × 600 pixel area with a speed of 2 pixels/frame. One trial comprised 150 frames. The starting points and trajectories of the dots were chosen randomly for each participant and each trial. The only restriction was that the distance between the dots had to be greater than 25 pixels at any time during movement. In the separated conditions, one trajectory for each dot was chosen randomly per trial, whereas in the grouped conditions, one trajectory common for all dots per trial was computed. The background was black during the whole experiment.

Stimulus generation and experimental control was accomplished with MATLAB 7, using the Psychophysics Toolbox 2.54 (Brainard, 1997).

Procedure. At the start of the experiment, each participant was shown the following instructions (in German) on the screen:

> Dear participant! Thank you for your readiness to participate in this experiment.
>
> Please imagine that you are working in the control center of an atomic power plant. A visual display that reflects the temperature of the cooling water is located on your control desk. Unfortunately, the computer that normally evaluates whether the temperature reflected in the visual display has crossed a critical limit or not is broken. Your task in this experiment is to replace this computer.
>
> For that purpose, the visual display is presented initially. Please look at it carefully. Then you are asked to guess whether the temperature has crossed the critical limit or not. Perhaps you will not be perfectly sure with your decision; therefore, a 7-point answer scale is available, ranging from 1 (*definitely not*) to 7 (*definitely*). At the beginning, you will have to guess. After a while, however, the decision will become easier as, after you deliver your rating, you will receive feedback whether the temperature had actually crossed the critical limit or not. For all of your answers, accuracy, rather than speed, is essential. Please do not take notes during the experiment. If you have any more questions, please ask them now. If you don't have any questions, please start the experiment by clicking on the "next" button.

When a participant asked a question, it was answered by the experimenter. When the participant clicked on the "next" button, the learning stage began with the first CS. Then the dots disappeared, and the question "Has the temperature crossed the critical limit?" was shown, along with a 7-point rating scale, with the anchors marked *definitely not* and *definitely*. After the participants marked one of the check boxes, they received feedback indicating either that the temperature was above the critical limit (in reinforced trials) or that it was below the critical limit (in nonreinforced trials). In the test stage, this feedback was omitted. Before the test stage, additional instructions were shown:

> In the following part of the experiment, we would like to survey what you have learned. For that purpose, there will be additional, new visual displays, but no feedback. Please look carefully at the visual display and enter your rating.

The training stage consisted of 15 reinforced and 15 nonreinforced trials. For the two added groups, this was AB+ and QR−; in the case of the two removed groups, it was ABCDE+ and QRSTU− (with different letters indicating different-colored dots and +/− indicating reinforced and nonreinforced trials). In the test stage, five trials of AB, ABCDE, QR, and QRSTU were displayed in the added groups. The removed groups were tested with AB, BC, CD, DE, AE, QR, RS, ST, TU, QU, and five trials of ABCDE and QRSTU each. Due to an experimental error, the first 6 participants did not receive the stimuli related to the nonreinforced trained stimulus in the test stage. However, since the training was the same, their test trials related to the reinforced trained stimuli were included in the analyses. All trial sequences were randomized but were kept parallel for the separated and the grouped conditions.

Results

For all statistical analyses, an α level of .05 and the Huynh–Feldt correction for the degrees of freedom was used.

Figure 1A shows that the rating of the reinforced and the nonreinforced stimuli diverged during training. In the last training trials, the mean ratings were 6.68 (SD = 0.657) and 1.26 (SD = 0.603). The training was analyzed with a 2 × 15 × 2 × 2 ANOVA that included the within-participants factors contingency (reinforced vs. nonreinforced) and trial (1–15) and the between-participants factors task (added vs. removed) and motion (separated vs. grouped). As expected, the main effect of contingency [$F(14,92) = 271.61$, $p < .001$, $\eta_p^2 = .75$], the main effect of trial [$F(14,1288) = 4.05$, $p < .001$, $\eta_p^2 = .04$], and the contingency × trial interaction [$F(14,1288) = 36.07$, $p < .001$, $\eta_p^2 = .28$] were significant. All other Fs were below

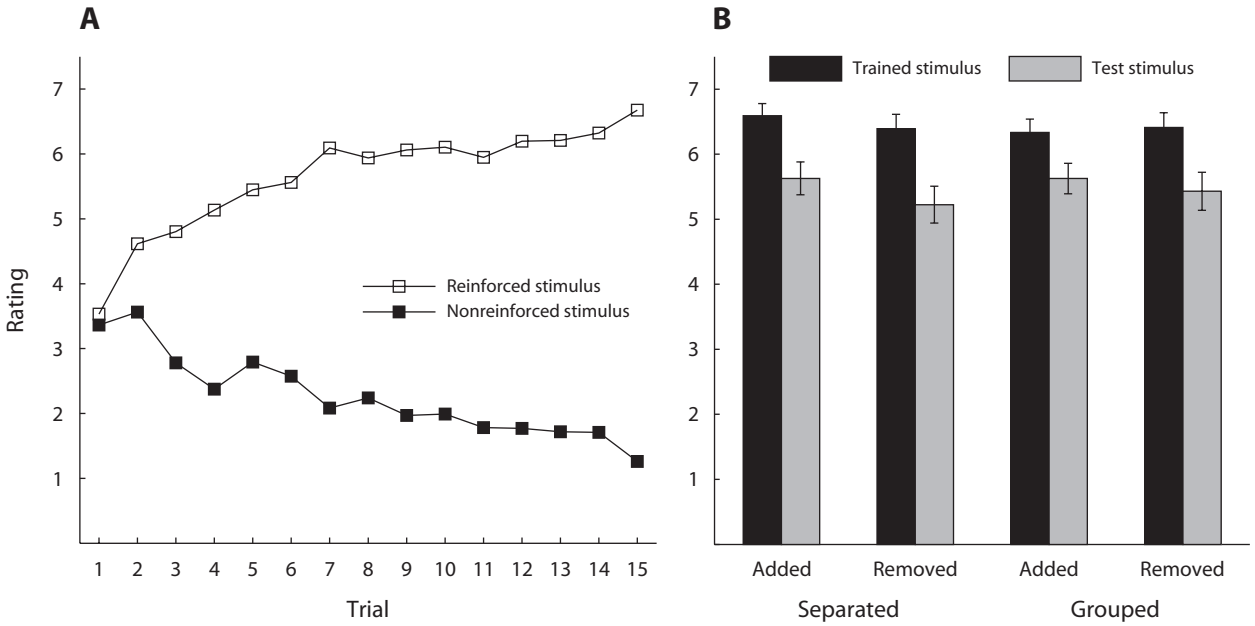


Figure 1. Mean ratings from 1 (*definitely not*) to 7 (*definitely*) in Experiment 1. Panel A shows ratings in the training stage, averaged across groups for the reinforced (AB+/ABCDE+) and nonreinforced (QR−/QRSTU−) stimulus compounds. Panel B displays, separately for the four experimental groups, the mean ratings of the reinforced trained stimulus and the corresponding test stimulus in the test stage.

1.70 (all $ps > .07$, all $\eta_p^2 s < .02$). Thus, independent of the task or the kind of motion, the discrimination was learned successfully.

Figure 1B displays the mean ratings in the test stage for the stimuli related to the reinforced trained stimulus. The black bars represent the rating of the previously trained stimulus itself, whereas the gray bars represent the rating of the corresponding test stimulus. In the added groups, for example, the rating of AB is shown by the black bar and the rating of ABCDE by the gray bar. The test stimulus was rated lower than the reinforced CS in all groups, and the decrement looks similar for all groups. To analyze the test stage, a $2 \times 5 \times 2 \times 2$ ANOVA was computed, including the within-participants factors cue (trained vs. test) and trial (1–5) and the between-participants factors task (added vs. removed) and motion (separated vs. grouped). Only the main effect of cue was significant [$F(1,92) = 67.00$, $p < .001$, $\eta_p^2 = .43$] (all other Fs < 1.47, $ps > .21$, $\eta_p^2 s < .02$). Thus, there was a generalization decrement in all groups. Neither the kind of task nor the kind of motion had any influence on the size of generalization decrements. Comparing the generalization tasks, η_p^2 for the main effect of cue was .47 in the removed groups and .39 in the added groups.

Although the nonreinforced stimuli were introduced to the experimental design only as filler cues (because the theories do not differ in their predictions), their results are reported for the sake of completeness: The mean rating for the nonreinforced trained stimulus (QR− for the added groups and QRSTU− for the removed groups) was 1.66 ($SD = 1.27$). The mean rating for the corresponding test stimulus was 2.7 ($SD = 1.44$). Thus, the newly introduced test stimuli were rated higher than the nonreinforced trained stimuli. We conducted a $2 \times 5 \times 2 \times 2$ ANOVA with the within-participants factors cue (trained vs. test) and trial (1–5) and the between-participants factors task (added vs. removed) and motion (separated vs. grouped). There was a significant main effect of cue [$F(1,86) = 44.02$, $p < .001$, $\eta_p^2 = .339$] and a significant cue × motion × trial interaction [$F(1,86) = 4.22$, $p < .05$, $\eta_p^2 = .047$]. All other effects were not significant (Fs < 2.25, $ps > .06$, $\eta_p^2 s < .025$).

Discussion

About two thirds of the participants successfully mastered the discrimination. Analyzing only "learners" ensured that the reinforced stimuli were rated high enough to provide the opportunity for observing decrements in the test. (Wheeler et al., 2006, have shown that failing to observe a decrement at all is often due to effects that are too small.) The test stimuli were rated lower than the reinforced training stimulus but above neutral; therefore, a generalization decrement was observed in all four groups. Comparing the added and removed groups, all decrements were of the same size. This symmetry is in line only with Pearce's configural model. Neither the REM nor Harris's (2006) model can account for these results. Thus, we have to conclude that the stimuli were processed configurally in both perceptual grouping conditions. The decrements were already symmetrical in the separated conditions; an additional grouping effect of motion in the grouped conditions was unobservable due to a floor effect.

Therefore, it is impossible to draw conclusions about the relationship of perceptual grouping and configural processing from the results of Experiment 1. Something must have influenced stimulus processing in Experiment 1,

because previous studies have consistently reported asymmetrical decrements (Brandon et al., 2000; Glautier, 2002; González et al., 2003; Wheeler et al., 2006). Although these studies differed in paradigms, stimuli, species, and experimental settings, they always observed asymmetry of generalization decrement. Therefore, one would expect that asymmetrical decrements are a quite robust phenomenon. Thus, the question arises: What distinguishes Experiment 1 from all these studies?

EXPERIMENT 2

A first difference between our Experiment 1 findings and those of other studies might be related to the motion of the stimuli. Normally, stimuli are presented at fixed positions during one trial. Motion per se may have had a configural impact in Experiment 1. Palmer (1999) described the new principle of synchrony as being *related* to the classical principle of common fate, in the sense that it is a dynamic factor, but *different* in that the "fate" of the elements does not have to be common, as long as the change occurs at the same time. Thus, the common feature of motion might have been enough to perceptually group the dots, even if each dot moved in a different direction. In the second experiment, we therefore "froze" the colored dots on their starting positions; the dots were presented at the starting positions, without movement.

Method

The materials and procedure in Experiment 2 were identical to those in Experiment 1, with the following exceptions.

Participants. We recruited 60 participants from the Philipps-Universität Marburg campus by posted advertisement and by face-to-face contact. We excluded 15 participants from further analysis because they did not solve the discrimination at the end of the training. The remaining 45 participants were from 18 to 39 years of age (M = 22.93). Thirty-four women took part. Twenty-one participants were tested in the added group, and 24 were tested in the removed group.

Stimuli. The different-colored dots were the same as those used in Experiment 1, but there was no motion. The position of the dots in the 600 × 600 pixel area in the center of the screen was chosen randomly for each participant and each trial. There was a minimal distance of 25 pixels between adjacent dots. The dots were shown for 150 frames, as before, but did not move.

Procedure. The training stage comprised 15 reinforced and 15 nonreinforced trials. The number of removed and added dots was reduced to 1, since the effects in Experiment 1 were large and the discrimination should be easier with three, instead of five, dots. The added group had to learn AB+ versus QR−, and the removed group had to learn ABC+ versus QRS−. Again, the trial sequences were randomized for each participant. In the test stage, AB, ABC, QR, and QRS were presented five times each in the added group, and the removed group was tested with AB, BC, AC, AB, AC, QR, RS, QS, QR, RS, and five trials of ABC and QRS each.

Results and Discussion

Figure 2A shows the mean rating during training. In the last trial, the mean rating for the reinforced stimulus was 6.84 (SD = 0.475) and, for the nonreinforced stimulus, 1.11 (SD = 0.383). The training scores were analyzed with a 2 × 15 × 2 ANOVA with the within-participants factors contingency (reinforced vs. nonreinforced) and trial (1–15) and the between-participants factor task (added vs. removed). The main effect of contingency [$F(14,43) = 603.99$, $p < .001$, $\eta_p^2 = .94$] and the main effect of trial [$F(14,602) = 2.38$, $p < .02$, $\eta_p^2 = .05$] were significant. Furthermore, the contingency × trial interaction [$F(14,602) = 48.49$, $p < .001$, $\eta_p^2 = .53$] and the triple interaction contingency × trial × task [$F(14,602) = 2.06$, $p < .05$, $\eta_p^2 = .05$] were significant (all other Fs < 1). The significant triple interac-

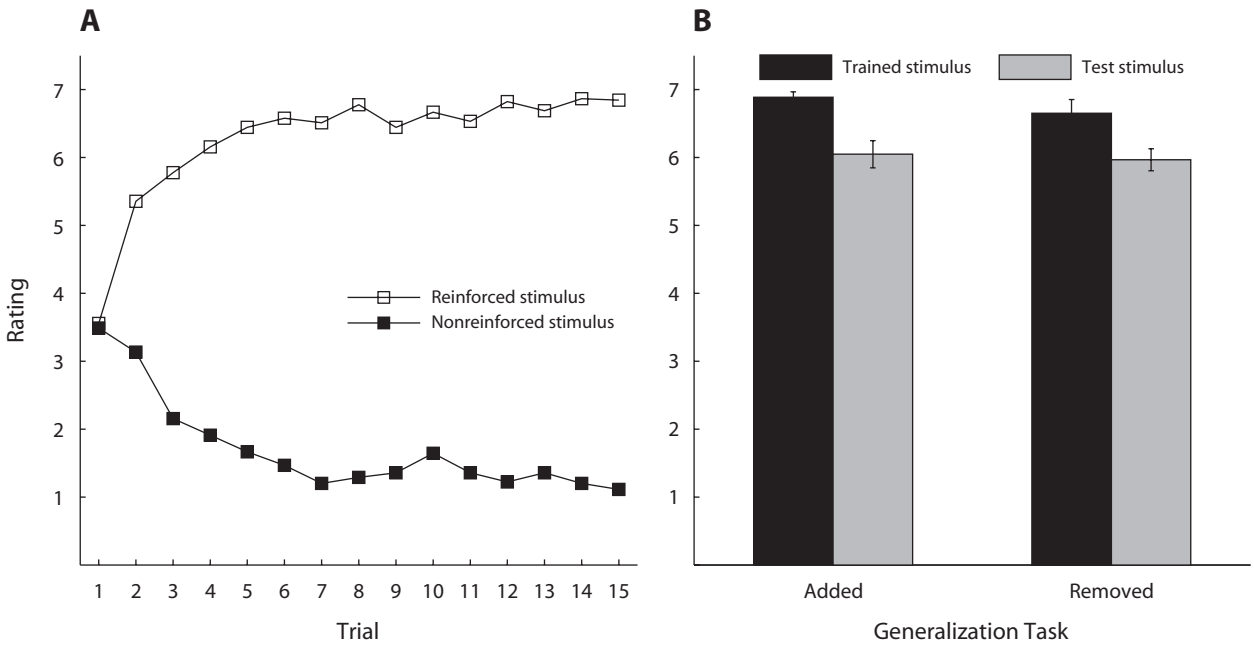


Figure 2. Mean ratings from 1 (*definitely not*) to 7 (*definitely*) in Experiment 2. Panel A shows ratings in the training stage, averaged across groups for the reinforced (AB+/ABC+) and nonreinforced (QR−/QRS−) stimulus compounds. Panel B displays, separately for two "generalization tasks," the mean ratings of the reinforced trained stimulus and the corresponding test stimulus in the test stage.

tion was probably due to differences in the first trial. On average, the added group rated the nonreinforced stimulus higher than the reinforced stimulus, whereas the removed group showed the reverse.

Figure 2B shows the mean ratings in the test stage of Experiment 2. For the removed group, the trained stimulus was ABC, and the test stimuli were AB, BC, and AC. For the added group, the trained stimulus was AB, and the test stimulus was ABC. The rating of the test stimuli was smaller than the rating of the trained stimuli in both groups.

To analyze the test stage, a 2 × 5 × 2 ANOVA was computed with the within-participants factors cue (trained vs. test) and trial (1–5) and the between-participants factor task (added vs. removed). Only the main effect of cue was significant [$F(1,43) = 33.06$, $p < .001$, $\eta_p^2 = .44$] (all other Fs < 1.23, ps $> .28$, η_p^2s $< .03$). Thus, there was a generalization decrement in all groups. The kind of task had no influence on the amount of generalization decrement. Comparing the main effect of cue in the added and removed tasks, η_p^2 was .46 for the removed group and .42 for the added group.

The mean rating for the nonreinforced trained stimulus (QR− for the added groups and QRS− for the removed groups) was 1.28 ($SD = 0.83$). The mean rating for the corresponding test stimulus was 1.96 ($SD = 0.89$). Thus, the newly introduced test stimuli were rated higher than the nonreinforced trained stimuli. We conducted an ANOVA with the within-participants factors cue (trained vs. test) and trial (1–5) and the between-participants factors task (added vs. removed). There was a significant main effect of cue [$F(1,43) = 31.25$, $p < .001$, $\eta_p^2 = .421$]. All other effects were not significant (Fs < 2.19, ps $> .12$, η_p^2s $< .048$).

The discrimination was mastered by about 75% of the participants. It is clear that it was not the motion of the dots that induced configural processing in Experiment 1, because the results of Experiment 2 were quite similar. This replicates the results of Experiment 1 but is still at variance with the previous studies. In order to enhance the comparability of the present study with these previous studies, we presented the stimuli on a fixed position in the middle of the screen in a third experiment.

EXPERIMENT 3

In contrast to previous studies, there were no fixed positions for the stimulus components in Experiment 2. Instead, the dots were randomly distributed across more or less the entire screen. This may have enhanced the impression of the whole screen as one stimulus and, thus, the different dots might have been processed as a whole. Therefore, the following experiment included a between-participants factor of position. In the *variable* group, the dots were randomly positioned as in Experiment 2; in the *fixed* group, the dots were presented on three selected positions in the middle of the screen. Furthermore, Experiment 2 revealed that adding and removing one dot to a two-dot compound was enough to generate stable decrements. This allowed us to use a within-participants design for the generalization task. All participants now received four stimulus compounds during training—two compounds with two dots and two with three dots, one reinforced each—and eight compounds in test—four with two and four with three dots (see Table 1). This resulted in 12 different colors needed for the whole experiment, instead of the 20 different colors that would have been needed in a case of within-participants design and four five-dot compounds in Experiment 1.

Method

The materials and procedure in Experiment 3 were the same as those used in Experiment 2, with the following exceptions.

Participants. Forty-one of 63 participants successfully solved the discrimination during training and were included in the analyses. They were from 18 to 29 years of age ($M = 21.27$). Thirty-one women took part. Nineteen participants were tested with variable- and 22 with fixed-dot positions.

Stimuli. Two additional colors (RGB: 0, 0, 224; 200, 255, 200) were defined, resulting in 12 different colors. Stimulus generation in the variable group was the same as in Experiment 2. In the fixed group, the same number and colors of dots were used, but the dots were presented at three fixed positions during the whole experiment. The positions were chosen so that the distance between them was equal to the mean distance of the dots in the variable group and that the dots were presented on a horizontal line in the middle of the screen. The colored dots were randomly assigned to a position in each trial.

Procedure. There were some slight modifications of the instructions, for clarification. Unlike in Experiments 1 and 2, the factor *task* was manipulated within participants. The training stage comprised 15 trials of AB+, CD−, EFG+, and HIJ−. In test, AB, CD, EFG, HIJ, ABX, CDY, EF, and IJ were presented twice each.

Results and Discussion

Figure 3A shows the rating during training. In the last trial, the mean ratings for AB+, CD−, EFG+, and HIJ− were 6.88 ($SD = 0.400$), 1.10 ($SD = 0.374$), 6.88 ($SD = 0.400$), and 1.12 ($SD = 0.400$), respectively. The training stage was analyzed with a 2 × 2 × 15 × 2 ANOVA with the within-participants factors contingency (reinforced vs. nonreinforced), task (added vs. removed), and trial (1–15) and the between-participants factor group (variable vs. fixed). The main effect of contingency was significant [$F(14,39) = 334.50$, $p < .001$, $\eta_p^2 = .90$], as was the contingency × trial interaction [$F(14,546) = 40$, $p < .001$, $\eta_p^2 = .51$] (all other Fs < 3, ps $> .10$, η_p^2s $< .07$).

Figure 3B shows the mean ratings in the test stage of Experiment 3. The ratings of the test stimuli ABX and EF were smaller than the ratings of the trained stimuli AB and EFG. To analyze the test stage, a 2 × 2 × 5 × 2 ANOVA was computed with the within-participants factors cue (trained vs. test), task (added vs. removed), and trial (1 or 2) and the between-participants factor group (variable vs. fixed). Only the main effect of cue was significant [$F(1,92) = 36.89$, $p < .001$, $\eta_p^2 = .49$] (all other Fs < 2.78, ps $> .10$, η_p^2s $< .07$). Thus, there was a generalization decrement in all groups. Neither the kind of task nor the kind of spatial arrangement had an influence on the amount of generalization decrement. Comparing the size of the main effects of cue in the added and the removed tasks, η_p^2 was .329 for AB versus ABX and .437 for EFG versus EF.

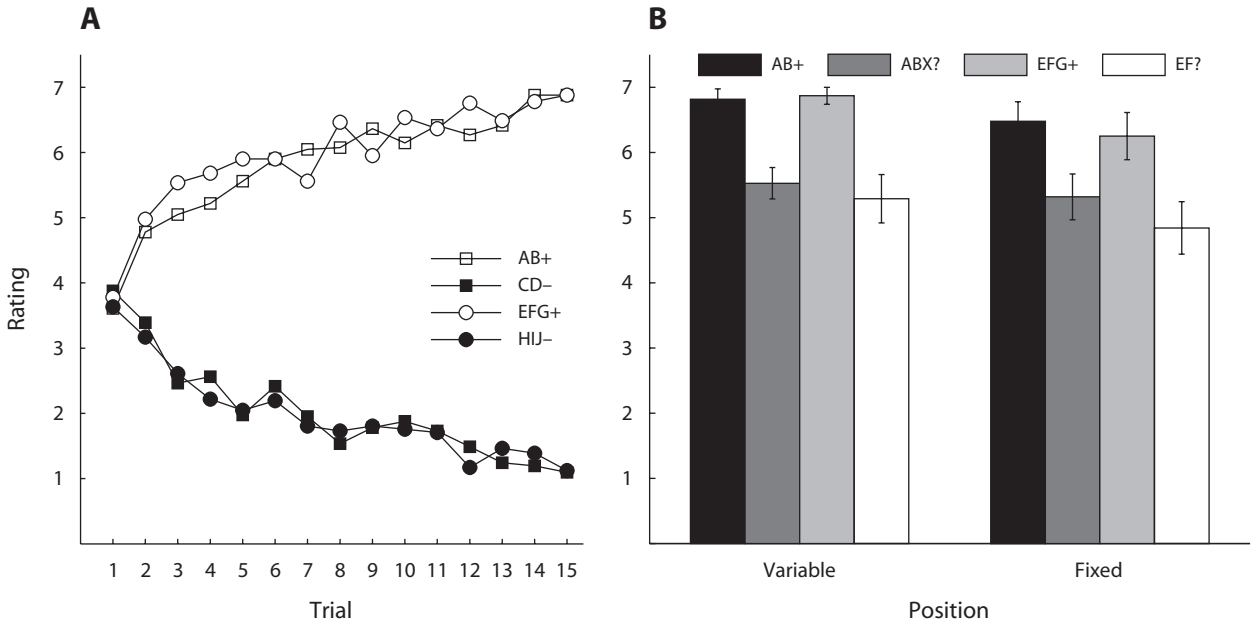


Figure 3. Mean ratings from 1 (*definitely not*) to 7 (*definitely*) in Experiment 3. Panel A shows ratings in the training stage, averaged across groups for the two reinforced and two nonreinforced stimulus compounds. Panel B displays, separately for the manipulation of the stimulus position, the mean ratings of the reinforced trained stimuli (AB+/EFG+) and the corresponding test stimuli (ABX?/EF?) in the test stage.

The mean ratings for CD, HIJ, CDY, and HI were 1.34 (SD = 0.83), 1.41 (SD = 1.02), 3.11 (SD = 1.77), and 2.96 (SD = 1.79), respectively. Thus, the newly introduced test stimuli were rated higher than the nonreinforced trained stimuli. We conducted an ANOVA with the within-participants factors cue (trained vs. test), task (added vs. removed), and trial (1 or 2) and the between-participants factor group (fixed vs. variable). There was a significant main effect of cue [$F(1,39) = 55.76$, $p < .001$, $\eta_p^2 = .588$] and a significant four-way interaction [$F(1,39) = 4.11$, $p < .05$, $\eta_p^2 = .095$]. All other effects were not significant (Fs < 3.07, ps $> .09$, $\eta_p^2 < .07$).

The results of Experiment 3 replicated the results of Experiments 1 and 2. About two thirds of the participants mastered the discrimination, and we found a stable decrement for the added and the removed generalization task in both groups. All decrements were of about the same size. This means that the presentation of the dots on fixed positions in the middle of the screen did not help to produce asymmetrical decrements, as observed in previous experiments in other labs.

GENERAL DISCUSSION

We compared generalization decrement in two different generalization tasks in three experiments. In the added task, a new stimulus was added to the previously trained CS, whereas in the removed task, a stimulus was removed from the previously trained CS. Experiments 1, 2, and 3, all using colored dots in a power-plant scenario (but different kinds of stimulus arrangements), resulted in symmetrical generalization decrements. This is the first time that symmetrical decrements have been reliably observed in studies investigating generalization decrements. This symmetry is predicted by Pearce's (1994) configural model, whereas Wagner's (2003) and Harris's (2006) elemental models both predict asymmetrical generalization decrements.

Before discussing potential factors controlling stimulus processing, and before explaining the divergence from the results of previous studies, we have to exclude an alternative explanation for the observed symmetrical decrements. In each of the discriminations of Experiments 1, 2, and 3, there was no need to rely on all the components presented as stimuli. Theoretically, the discriminations could have been solved by using only one component (one colored dot). If indeed the participants had relied on only one component (for example, the most salient one), removing "irrelevant" components would have led to no decrement at all. On the other hand, the decrement after removing the "relevant" component should have been at maximum, since the stimulus that carried all the associative strength was no longer present at test. Therefore, elemental models still predict that removing a stimulus should result in more of a decrement on average than adding a stimulus, even if only part of the stimulus compound is used to solve the discrimination. Nevertheless, future studies of generalization decrements should address this issue more directly—for example, by assessing whether all single stimuli are rated alike.

The results of Experiments 1, 2, and 3 suggest that the stimuli were processed configurally and that this processing was not influenced by the kind of motion, the motion per se, or the spatial arrangement of the stimuli, although other studies have shown an impact of the kind of stimulus presentation (e.g., Glautier, 2002; Livesey & Boakes,

2004). Furthermore, the spatial arrangement in the fixed group of Experiment 3 was similar to that of Wheeler et al.'s (2006) Experiment 2, and both experiments used clearly separable visual stimuli. But for all of that, the results differ. What in the power-plant experiments was so powerful in inducing configural processing that omitting other potential manipulations, such as the common motion of the dots, did not result in elemental processing? Three factors come to mind in comparing these three experiments with previous studies—especially the causal learning studies of Wheeler et al. and Glautier (2004).

A first hypothesis concerns within-trial timing. From other studies in our lab comparing causal learning and classical conditioning, we have clues that the amount and pressure of time may influence stimulus processing. Lachnit, Schultheis, König, Üngör, and Melchers (2008) and Lachnit, Thorwart, and Schultheis (2008) observed different kinds of stimulus processing in different paradigms and response systems. They used modified positive and negative patterning designs with two different kinds of visual stimuli (pictures of foods, colored rectangles) in eyelid-conditioning, skin-conductance-conditioning, and causal-learning experiments. The kind of stimuli had no impact on stimulus processing. Instead, systematic differences were observed, depending on the response system and the time for stimulus processing. Pearce's configural model was best in predicting early stimulus processing in skin conductance conditioning, whereas elemental models were superior in causal-learning studies in which stimulus duration was paced by the participants. In conditioning experiments with fixed interstimulus intervals (ISIs), there is a time limitation, and further results provided some hints that reducing the ISI (and thus enhancing time pressure) in an eyelid-conditioning study influenced learning toward configural processing. In sum, Pearce's configural model predicted early stimulus processing; other theories predicted late stimulus processing better. This factor might explain the results of our present experiments. In the power-plant studies, stimulus presentation was limited to 2 sec, whereas the ISI in Wheeler et al.'s (2006) experiments was paced by the participants. Thus, in human contingency learning with time pressure, correct predictions are made by Pearce's configural theory and, without time pressure, the predictions of elemental theories are correct. Hence, reducing the time pressure in the power-plant scenario should result in asymmetrical decrements. Glautier's (2004) results contradict this explanation: His planes traversed the screen in approximately 2.1 to 2.8 sec, and he found no decrement for the added stimuli. However, Wheeler et al. have already suggested alternative explanations for Glautier's (2004) results. Furthermore, since there was no need to respond during training in Glautier's paradigm, the issue of time pressure on stimulus processing may not have been relevant for learning.

Second, there is a further difference among the scenarios used in the different studies. Wheeler et al.'s (2006) food-poisoning scenario, as well as Glautier's (2004) air pollution scenarios, are causal scenarios. The participants had to rate the likelihood that the meals would cause food poisoning or how much pollution was caused by an airplane, respectively. In the power-plant scenario of the present experiments, no statement was made about the causal relationship between the dots and the cooling water's temperature. Instead, it was only said that there is a visual display that reflects the temperature of the cooling water and that the participants should predict whether the temperature had crossed the critical limit. De Houwer, Beckers, and Glautier (2002) have shown that blocking was stronger when the cues were presented as potential causes of the outcome rather than as mere indicators or predictors of the outcome. In the causal conditions of De Houwer et al.'s experiments, the participants were asked to rate the likelihood that firing a weapon would be followed by the destruction of a tank (i.e., the outcome), whereas in the predictive conditions, they were asked to rate the likelihood of the destruction of the tank when an abstract visual figure was present. Pineño, Denniston, Beckers, Matute, and Miller (2005) observed reliable overshadowing only between causes and only when the test question was causal. In order to study and contrast predictive and causal learning, they either used different wording in instructions about the relationship of two different stimuli and an outcome or took advantage of participants' preexperimental knowledge. Furthermore, test questions asked for the predictive and causal relationships separately. Therefore, the symmetry in our experiments may not have been due to a large decrement in the added groups but to reduced overshadowing (i.e., small decrements) in the removed groups in predictive contingency learning tasks. It requires the use of a causal scenario with moving dots to test this hypothesis.

Finally, it may be possible to provide an elemental explanation for the symmetrical decrements obtained here, in terms of generalization among stimulus components. Our discrimination tasks involved a large number of colors and, perhaps, there was generalization between single components (e.g., from a red dot to an orange dot). Furthermore, since the nonreinforced stimuli were rated well below the neutral level, it seems reasonable to assume generalization not only of excitation, but also of inhibition. If C, D, or E are similar in color to the nonreinforced Q or R, this would reduce the prediction given for ABCDE after training AB+. Thus, what has been interpreted as external inhibition could have been generalization of inhibition from the nonreinforced cues Q and R. Of course, there is no reason to expect a systematic generalization effect of this sort; colors were randomly allocated to conditions. However, if it occurred only in some proportion of participants, this may have been sufficient to produce the observed amount of response decrement to ABCDE. And although it is true that one would expect an equal similarity relationship between C, D, and E to A and B, this might not produce complementary evidence for excitatory generalization for two reasons: First, ratings of the AB compounds were already at ceiling values (close to 7), which would have prevented any further increment in ratings by the participants for whom C, D, or E were similar to A and B. Second, the presence of A and B colors within the same test stimulus would presumably reduce generalization, because participants would more readily recognize that colors C, D, and E were not the same as A

or B. As stated here, this argument relies on generalization from Q and R to C, D, or E, due to a possible similarity in hue. A way to control this would be to use stimulus material that is less likely to support generalization between the components, as could be the case with labeled pictures of food (as Wheeler et al., 2006, used in their experiments) or, more relevant to the present study, with labels of colors.

AUTHOR NOTE

This research is part of the doctoral thesis of A.T. and was supported by the Deutsche Forschungsgemeinschaft graduate program NeuroAct (DFG 885). H.L. was supported by Grant La 564/19-1. We thank Steven Glautier for supporting A.T. in the course of her research visit to his lab in Southampton. Correspondence concerning this article should be addressed to A. Thorwart, Department of Psychology, Philipps-Universität Marburg, Gutenbergstraße 18, 35032 Marburg, Germany (e-mail: anna.thorwart@staff.uni-marburg.de).

REFERENCES

BRAINARD, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, **10**, 433-436.

BRANDON, S. E., VOGEL, E. H., & WAGNER, A. R. (2000). A componential view of configural cues in generalization and discrimination in Pavlovian conditioning. *Behavioural Brain Research*, **110**, 67-72.

DE HOUWER, J., BECKERS, T., & GLAUTIER, S. (2002). Outcome and cue properties modulate blocking. *Quarterly Journal of Experimental Psychology*, **55A**, 965-985.

GLAUTIER, S. (2002). Spatial separation of target and competitor cues enhances blocking of human causality judgements. *Quarterly Journal of Experimental Psychology*, **55B**, 121-135.

GLAUTIER, S. (2004). Asymmetry of generalization decrement in causal learning. *Quarterly Journal of Experimental Psychology*, **57B**, 315-329.

GLAUTIER, S. (2008). The elemental–configural distinction: A problem of two dimensions: Commentary on Melchers, Shanks, and Lachnit (2007). *Behavioural Processes*, **77**, 431-433.

GONZÁLEZ, F., QUINN, J. J., & FANSELOW, M. S. (2003). Differential effects of adding and removing components of a context on the generalization of conditional freezing. *Journal of Experimental Psychology: Animal Behavior Processes*, **29**, 78-83.

HARRIS, J. A. (2006). Elemental representations of stimuli in associative learning. *Psychological Review*, **113**, 584-605.

LACHNIT, H., SCHULTHEIS, H., KÖNIG, S., ÜNGÖR, M., & MELCHERS, K. G. (2008). Comparing elemental and configural associative theories in human causal learning: A case for attention. *Journal of Experimental Psychology: Animal Behavior Processes*, **34**, 303-313.

LACHNIT, H., THORWART, A., & SCHULTHEIS, H. (2008). *Comparing elemental and configural associative theories in human Pavlovian conditioning with indicators of early and late CS processing*. Unpublished manuscript.

LILJEHOLM, M., & BALLEINE, B. (2008). It's elemental my dear Watson. *Behavioural Processes*, **77**, 434-436.

LIVESEY, E. J., & BOAKES, R. A. (2004). Outcome additivity, elemental processing and blocking in human causality judgements. *Quarterly Journal of Experimental Psychology*, **57B**, 361-379.

LIVESEY, E. J., & HARRIS, J. A. (2008). What are flexible representations? Commentary on Melchers, Shanks and Lachnit. *Behavioural Processes*, **77**, 437-439.

MELCHERS, K. G., SHANKS, D. R., & LACHNIT, H. (2008). Stimulus coding in human associative learning: Flexible representations of parts and wholes. *Behavioural Processes*, **77**, 413-427.

PALMER, S. E. (1999). *Vision science: Photons to phenomenology*. Cambridge, MA: MIT Press.

PAVLOV, I. P. (1927). *Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex* (G. V. Anrep, Trans. & Ed.). London: Oxford University Press.

PEARCE, J. M. (1987). A model for stimulus generalization in Pavlovian conditioning. *Psychological Review*, **94**, 61-73.

PEARCE, J. M. (1994). Similarity and discrimination: A selective review and a connectionist model. *Psychological Review*, **101**, 587-607.

PINEÑO, O., DENNISTON, J. C., BECKERS, T., MATUTE, H., & MILLER, R. R. (2005). Contrasting predictive and causal values of predictors and of causes. *Learning & Behavior*, **33**, 184-196.

SCHULTHEIS, H., THORWART, A., & LACHNIT, H. (2008a). HMS: A MATLAB simulator of the Harris model of associative learning. *Behavior Research Methods*, **40**, 442-449.

SCHULTHEIS, H., THORWART, A., & LACHNIT, H. (2008b). Rapid-REM: A MATLAB simulator of the replaced-elements model. *Behavior Research Methods*, **40**, 435-441.

SHANKS, D. R., LACHNIT, H., & MELCHERS, K. G. (2008). Representational flexibility and the challenge to elemental theories of learning: Response to commentaries. *Behavioural Processes*, **77**, 451-453.

THORWART, A., SCHULTHEIS, H., KÖNIG, S., & LACHNIT, H. (in press). ALTSim: A MATLAB simulator for current associative learning theories. *Behavior Research Methods*.

WAGNER, A. R. (2003). Context-sensitive elemental theory. *Quarterly Journal of Experimental Psychology*, **56B**, 7-29.

WHEELER, D. S., AMUNDSON, J. C., & MILLER, R. R. (2006). Generalization decrement in human contingency learning. *Quarterly Journal of Experimental Psychology*, **59**, 1212-1223.

NOTES

1. To avoid misunderstandings, the term *component* is used for distinct parts of a stimulus compound; for example, a stimulus compound comprises its two components—a red light and a yellow light. *Element* refers to the even smaller parts of a component's representation, as hypothesized by the elemental models.

2. The REM is sometimes labeled as a "hybrid" model (Glautier, 2008; Liljeholm & Balleine, 2008; Livesey & Harris, 2008). However, because learning still occurs about elements and not about whole configurations, we classify the REM as a purely elemental theory. In contrast, Pearce's (1994) model assumes that the stimulus configuration as a whole gets associated. It is this kind of conditioning we refer to as *configural processing* in the present article.

(Manuscript received July 25, 2008;
revision accepted for publication September 11, 2008.)

Running head: GENERALIZATION DECREMENTS AND SALIENCE

Generalization decrements and salience: new data and more elaborated analyses

Anna Thorwart* and Harald Lachnit

Philipps-Universität Marburg, Germany

*Corresponding author

Addresses: Anna Thorwart and Harald Lachnit
Fachbereich Psychologie
Philipps-Universität Marburg
Gutenbergstraße 18
35032 Marburg, Germany
Phone: +49 6421 2823715
Fax: +49 6421 2826621
Email: anna.thorwart@staff.uni-marburg.de

Abstract

Thorwart and Lachnit (2009) had found reliable symmetrical decrements in two generalization tasks: Removing an already trained component from a compound did not result in larger decrements than adding a new one. Surprisingly, the present experiments revealed asymmetrical decrements independent of the manipulations (duration of stimulus presentation, cover story) introduced. However, more precise analyses and simulations showed that the underlying patterns of responding in the present experiments and in Experiment 1 of Thorwart and Lachnit were similar. All results can be predicted by the configural model of Pearce (1994), if differences in the saliences of the components are allowed. Such differences became relevant if more and more complex stimuli had to be processed.

Generalization decrements and salience: new data and more elaborated analyses

Traditionally, many well known associative learning phenomena have been explained in a so-called elemental manner. Here it is assumed that stimulus components are represented as separate entities and that the overall associative strength of a compound (e.g., AB) is based on the algebraic sum of the associative strengths of its components A and B (e.g., Mackintosh, 1975; Pearce & Hall, 1980; Rescorla & Wagner, 1972). On the other hand, however, there is considerable evidence that humans as well as non-human animals can successfully handle discrimination problems that cannot be solved in such a purely elemental manner. Therefore, more recent elemental models like the Replaced Elements Model (REM; Wagner, 2003) and the model of Harris (2006), assume that a compound may not activate exactly the same entities and associations as its components and that there is some variability in that. Others have suggested instead that stimuli are always processed configurally. In this view, a compound AB consisting of stimuli A and B maps onto a representational entity, a so called configural unit, which is distinct from its elements (say Z). Only this configural unit enters into an association with the outcome. Furthermore, with respect to behavior it is assumed that generalization between such entities takes place based on their component similarity (e.g., Pearce, 1994).

Empirically, there are different possibilities to test these rivaling theories. Firstly, for certain discrimination problems these theories clearly differ in their predictions about which discriminations should be *acquired* more rapidly (see e.g., Harris, Livesey, Gharaei, & Westbrook, 2008; Kinder & Lachnit, 2003; Lachnit, Schultheis, König, Üngör, & Melchers, 2008; Redhead & Pearce, 1995; Thorwart, Glautier, & Lachnit, 2009). Secondly, these theories differ with respect to responding to new stimuli in a test after acquisition (e.g., summation: Myers, Vogel, Shin, & Wagner, 2001; Glautier, Redhead, Thorwart, & Lachnit, in press; Pearce, Aydin, & Redhead, 1997; generalization decrements: Brandon, Vogel, & Wagner, 2000; Glautier, 2004; González, Quinn, & Fanselow, 2003; Thorwart & Lachnit, 2009; Wheeler, Amundson, & Miller, 2006). Most current theories of associative learning assume that stimuli are invariably processed in one and the same way. As a

consequence, the question of whether stimuli are processed elementally or configurally is usually discussed in an either/or manner (see Pearce & Bouton, 2001; Wasserman & Miller, 1997, for recent reviews). One explanation of the mixed results observed so far, however, could be that there exists flexibility in the way in which humans and animals deal with discrimination problems (for details see Melchers, Shanks, & Lachnit, 2008).

In the light of this potential flexibility, we (Thorwart & Lachnit, 2009) investigated generalization decrements - where so far unequivocal support for elemental theories had been found – in order to determine factors that might influence the kind of processing there, too. We used two different generalization tasks introduced by Brandon, Vogel and Wagner (2000). In the Added Cue task, a new stimulus component is added to the previously trained stimulus (AB+→ ABX) whereas in the Removed Cue task, a stimulus component is removed from the previously trained stimulus (EFG+→ EF). The predictions of the rivaling models were simulated by using several MATLAB based programs (Schultheis, Thorwart, & Lachnit, 2008a, 2008b, Thorwart, Schultheis, König & Lachnit, 2009). Pearce's (1994) configural model predicts symmetrical decrements (that is the amount of decrement produced by testing ABX after training with AB will be the same as that produced in testing with EF after training with EFG). On the other hand, both elemental theories (Harris, 2006; Wagner, 2003) predict asymmetrical generalization decrements: removal of components always results in a larger decrement than adding. This prediction of asymmetry remains unaffected by the supposed variability of their stimulus representation.

Previously, only asymmetrical decrements had been reported in human (Glautier, 2004; Wheeler, Amundson & Miller, 2006) as well as in non-human learning experiments (Brandon, Vogel, & Wagner, 2000; González, Quinn, & Fanselow, 2003). Guided by the question "under which conditions is processing best described by configural theories, and under which conditions is it best described by elemental theories?", we attempted to investigate whether manipulation of the stimuli is able to produce symmetrical generalization decrements. In the first experiment of Thorwart

and Lachnit (2009) the Gestalt principle of common fate was used according to which objects moving with the same speed in the same direction become perceptually grouped and are seen as belonging together (Palmer, 1999). Five or two dots differing in color were presented moving across a black screen for two seconds either all in the same direction or in different directions. However, symmetrical decrements were observed in both conditions. In Experiment 2 of Thorwart and Lachnit (2009) we "frozen" the colored dots on their starting position (i.e. the dots were presented located at the starting positions without movement). In addition, the number of removed and added dots was reduced to one as the effects in Experiment 1 were large and the discrimination should be easier with three instead of five dots. Again, however, symmetrical generalization decrements were observed. Experiment 3 included a between-subjects factor of position: in the Group Variable, the dots were randomly positioned as in Experiment 2; in the Group Fixed, the dots were presented on three selected positions in the middle of the screen. Again, symmetrical decrements were observed under all conditions.

We discussed three potential reasons for this reliable divergence, two of which (duration of stimulus presentation, cover story of the task) we will focus on in the experiments of the present article. Firstly, from other studies in our lab comparing causal learning and classical conditioning, we have clues that the stimulus duration and thus the pressure of time may influence stimulus processing (Lachnit et al., 2009). In all three experiments, stimulus presentation had been limited to 2 sec whereas the stimulus duration for instance in Wheeler et al.'s (2006) causal learning experiments had been self-paced by the participants. Hence, reducing the time pressure by a prolonged stimulus presentation may result in asymmetrical decrements. Secondly, even though all mentioned human studies were contingency learning experiment, they used different cover stories. The food-poisoning scenario of Wheeler et al. (2006) as well as the air pollution scenario of Glautier (2004) is a causal scenario. The participants had to rate the likelihood that the meals cause food poisoning or how much pollution is caused by an airplane, respectively. In the power plant scenario of Thorwart and Lachnit (2009), no statement was made about the causal relationship

between the dots and the cooling water's temperature. Pineño, Denniston, Beckers, Matute, and Miller (2005) observed reliable overshadowing only between causes and only when the test question was causal. As the Removed Cue task just measures overshadowing, the symmetry in our experiments may have been due to reduced overshadowing (i.e., small decrements) in the Removed Cue task because of the predictive instruction.

Experiment 1

The present experiment explicitly manipulated stimulus duration between subjects and used a causal scenario in order to enhance the possibility of asymmetrical decrements. As in Experiment 1 of Thorwart and Lachnit (2009) the influence of perceptual grouping by motion on the stimulus processing was investigated. Again, the cues were small dots moving across a black screen. Color of the dots was the decisive feature to solve the discrimination whereas motion, number of dots, and duration were irrelevant for predicting the outcome. Thus, participants received four stimulus compounds during training: two compounds with two dots and two with three dots, one of each followed by the outcome (AB+, CD-, EFG+, HIJ-). In the test stage, a new dot was added to the two-dot-compounds (ABX, CDY) whereas one dot was removed from the three-dot-compounds (EF/EG/FG, HI/HJ/IJ). To evaluate the models' predictions, generalization decrements for the reinforced stimuli in the Removed Cue and Added Cue task were compared within-subject. A between-subject factor controlled whether the dots were presented as long as in Thorwart and Lachnit (2 sec) or four times as long (8 sec). The kind of motion varied as a second between-subjects factor: in the Separated Motion conditions, the direction of motion was randomly chosen for each dot and trial (with some restrictions); in the Grouped Motion conditions, all dots were moving in one and the same direction (chosen randomly for each trial). In Experiment 1 of Thorwart and Lachnit (2009), we had hypothesized that perceptually grouped stimuli should be processed according to configural principles whereas non-grouped stimuli are prone to elemental processing. As there already had been symmetrical decrements in the Separated Motion condition, any additional influence of the grouped motion could not have been

observed. In the present experiment, this floor effect should be avoided at least in the groups with the prolonged stimulus duration. Additionally, the instruction introduced a medical cover story with a causal link between the dots (bacteria in the blood) and the outcome (disease) to further increase the likelihood of asymmetrical decrements, i.e. elemental processing, and decrease the likelihood of a floor effect.

Method

Participants. We recruited 159 participants from the university campus by posted advertisements. They received course credits or were paid. Participants were included in further analysis if they had solved the discrimination at the end of the training. To be included, the rating in the last two trials of training had to be equal or higher than 5 for the reinforced stimuli and equal or lower than 4 for the non-reinforced stimuli (these numbers refer to scores on an 8-point scale; see below). The remaining 126 participants (30 men and 96 women) were aged from 19 to 40 years with a mean age of 22.53. In four groups - Separated-Short, Separated-Long, Grouped-Short, and Grouped-Long - 33, 33, 31, and 29 participants were tested.

Apparatus and Stimuli. The experiment took place in labs of the Department of Psychology. In the labs, stimuli were presented on 17-inch color monitors running in a 1280 (width) x 1024 (height)-pixel resolution and with a 75 Hz frame rate.

Different colored dots (17 pixels) were used as the cues. Twelve different colors were needed for the whole experiment (RGB: 64 204 255; 53 169 55; 234 50 24; 147 36 207; 238 29 90; 187 212 65; 231 95 217; 69 49 181; 255 157 0; 253 228 111; 0 0 224; 200 255 200). The colors were assigned randomly for each participant, but the Separated Motion and Grouped Motion groups were matched. The dots moved across a central 750 x 750 pixel area with a speed of one pixels/frame. One trial composed 150 frames in the Short and 600 frames in the Long Stimulus Duration condition, resulting in 2 and 8 sec, respectively. (For 13 participants of the Group Separated-Long the frame rate was 60 Hz, resulting in additionally prolonged stimulus duration of 10 sec. Analyses of this group included this factor revealed no significant influence ($F < 1.2$)). The starting points and trajectories of the dots were chosen randomly for each

participant and each trial. The only restriction was that the distance between the dots had to be greater than 25 pixels at any time during movement. In the Separated Motion conditions, one trajectory for each dot was chosen randomly per trial whereas in the Grouped Motion conditions, one trajectory common for all dots per trial was computed. The background was black during the whole experiment. Stimulus generation and experimental control was accomplished by MATLAB 7 using the Psychophysics Toolbox 2.54 (Brainard, 1997).

Procedure. At the start of the experiment, each participant read the following instructions (in German) on the screen:

> Dear participant! Thank you for your readiness to participate in this experiment.
>
> In the following experiment, you will learn which bacteria in the blood are causal for a new illness. Please imagine that you are working in a medical laboratory. Recently there were several cases of this new illness. In order to find the cause for this illness, you use a new procedure for investigating bacteria in blood samples.
>
> You will see the sample of a test person on the screen at a time. Afterwards please indicate whether the bacteria in this blood sample cause a disease, thus whether the test person has gotten sick or not. Perhaps you will not be perfectly sure with your decision, therefore an 8-point answer scale is available, ranging from "1 = definitely not ill" to "8 = definitely ill". After you delivered your rating, you will receive feedback whether the disease occurred or not. At the beginning you will have to guess but after a while you will learn what causes a disease. It doesn't matter how fast you answer, therefore, take the time you need. If you have any question, please ask now. Otherwise, please start the experiment by clicking on the "next" button.

Each trial started with a pattern of dots moving across the black screen. After the dots disappeared after either 2 or 8 sec, the question "Are the bacteria in this blood sample causing a disease?" was shown together with an 8-point rating scale with the anchors marked "definitely not ill" and "definitely ill". After the participants marked one of the check boxes, they received feedback indicating if the patient was ill or not. Before the test stage, a further instruction told the participants that during the next experimental phase there will be additional new stimuli and no feedback anymore.

The training stage comprised 15 trials of AB+, CD-, EFG+, and HIJ-, respectively, with each letter corresponding to one color. “+” indicates that the patient was ill, ”-“ that he was not. The test stage consisted of 36 test trials: three trials each of AB, CD, EFG, HIJ, ABX, CDY, respectively, and one trial each of EF, GH, EG, HI, IJ, HJ, A, B, C, D, E, F, G, H, I , J, X, Y. Thus, every dot of the trained stimulus EFG was removed once. Test trials consisting only of one dot were included for the sake of completeness. However, as the Grouped Motion versus Separated Motion manipulation cannot be implemented with only one dot, those trials cannot be compared with the other trials properly. Trial sequences were randomized, but kept parallel between groups.

Results and Discussion

For all statistical analyses, α was set to .05 and the Huynh-Feldt correction of the degree of freedoms was used where appropriate.

Figure 1 displays the mean rating for the relevant stimuli in the test stage. The upper panels show the results of both groups with prolonged stimulus duration of 8 sec, the lower panels for the groups with the previously used, shorter stimulus duration of 2 sec. The groups with dots moving in different directions (Separated Motion) are shown on the left side, groups with dots moving in the same direction (Grouped Motion) on the right side. Within each panel, the black bars represent the mean rating of the previously trained stimulus (AB for the Added Cue task on the left side and EFG for the Removed Cue task on the right side) whereas the white bars represent the rating of the corresponding test stimulus (ABX and EF/EG/FG,

respectively). First of all, the results did not differ between the groups. In all groups the ratings of test stimuli were lower than those of the reinforced training stimuli. Moreover, the decrement seemed to be larger for the Removed Cue than for the Added Cue task, independent of the manipulation of the stimulus duration or the motion. This was confirmed by a 2 x 2 x 3 x 2 x 2 ANOVA including the within-participants factors cue (trained vs. test), task (added vs. removed) and block (1-3), and the between-participants factors stimulus duration (2 sec vs. 8 sec) and motion (grouped vs. separated). There were main effects of cue (F (1, 122) = 205.225, p < .001, η_p^2= .627), task (F (1, 122) = 9.71, p < .003, η_p^2= .074) and motion (F (1, 122) = 7.75, p < .007, η_p^2= .06) as well as a Cue*Task interaction (F (1, 122) = 10.653, p < .002, η_p^2 = .08). The unexpected main effect of motion was due to higher ratings in the Grouped Motion condition (mean = 7.44) compared to the Separated Motion condition (mean = 7.18).

Fig. 1 about here

The present experiment revealed asymmetrical decrements: Removing an already trained component from a three-dot compound resulted in larger decrements than adding a new one to a two-dot compound. In our previous experiments, we had observed reliable symmetrical decrements in human predictive learning using the same kind of stimulus material and a stimulus duration of 2 sec. We had hypothesized that the symmetrical decrements observed by Thorwart and Lachnit (2009) had been due to time pressure during stimulus processing. However, stimulus duration in the present experiment had no effect on the decrements' symmetry. Decrements were asymmetrical with stimulus duration of 2 as well as of 8 sec. Therefore, time pressure is not sufficient to induce symmetrical decrements.

Consistent with Experiment 1 of Thorwart and Lachnit (2009), the kind of motion did not influence the decrements' (a)symmetry. As decrements in groups with separated motion were asymmetrical in the present experiment, there was no floor

effect here. Therefore, we have to conclude that the motion – grouped or separated - is also not sufficient for symmetrical decrements.

The simplest explanation for the differing results so far is that a causal cover story is necessary and sufficient for asymmetrical decrements. In Thorwart and Lachnit (2009), a power plant scenario without any information about the causal link between the cues and the outcome was used. In the present study, the instruction highlighted the causal relationship between the bacteria and the disease as outcome. This conclusion is supported by the results of previous causal learning studies of Wheeler et al. (2003) and Glautier (2004) which had observed asymmetrical decrements. Furthermore, this would be in line with the suggestion of Pineño et al. (2005) that there is greater susceptibility to cue competition in causal than in predictive attribution. In elemental models, the decrement in the Removed Cue task is due to overshadowing, i.e. cue competition during the training stage, whereas the decrement in the Added Cue task is based on changes in stimulus representation in the test stimuli. Following Pineño et al. (2005), changing the predictive cover story to a causal one should result in an increased decrement for the removed but not for the Added Cue task. In configural models, both decrements are due to a decreased activation of the previously trained configural unit. Changing the cover story should not have any differential effect. An influence of the cover story then implies that stimulus processing in the present experiment followed principles incorporated in elemental models. On the other hand, the differing susceptibility to cue competition is only based on the observation of Pineño et al. (2005) and, as they pointed out, cannot be explained by any contemporary – elemental or configural - model of associative learning in its current form. For example, Wagner's REM has a parameter accounting for perceptual interaction (r = proportion of context-dependent elements) that influence the amount of overshadowing but not the asymmetry of the decrements. Therefore, an influence of the cover story on the asymmetry would pose a major challenge to associative learning models. The second experiment looks for further empirical evidence.

Experiment 2

The following experiment investigated whether the causal cover story influenced the symmetry of the generalization decrements in Experiment 1. One group received the instruction of Experiment 1. These participants were told that they have to learn which bacteria in blood samples are causal for a new disease. The instruction of the Group Predictive stated that participants will learn which components of a blood sample are indicators for a new disease and allow to diagnose it. Furthermore, the rating question of the Group Causal asked "Are the bacteria in this blood sample causing the disease?" whereas the Group Predictive was asked "Are the components in this blood sample indicating a disease?".

As Experiment 1 replicated the missing influence of perceptual grouping by motion of Thorwart & Lachnit (2009) and revealed no effect of stimulus duration, the dots in the following experiment were displayed on fixed positions in the middle of the screen for 2.5 sec, similar to Experiment 3 of Thorwart & Lachnit.

Methods

Only differences to the previous experiment are reported.

Participants. 43 of 63 participants were included in the analyses as they successfully solved the discrimination during training. They were aged from 19 to 34 years with a mean age of 21.3. 13 men and 30 women took part. 21 participants were tested with predictive and 22 with the causal instruction.

Apparatus and Stimuli. The experiment took place either in labs in the Department of Psychology or in a room of the university's student services house. In the student services house, an IBM Thinkpad was used which was placed in a paperboard cabin. All monitors were 15-inch and running in a 1024 (width) x 768 (height)-pixel resolution and with a 60 Hz frame rate. The participants wore sound-attenuating ears muffs. The 16-pixel dots were presented on a horizontal line in the middle of the screen for 150 frames, resulting in 2.5 sec stimulus presentation. The colored dots were randomly assigned to a position in each trial.

Procedure. The following instruction was presented in the Group Predictive,

with changes to the causal instruction are highlighted:

> Dear participant! Thank you for your readiness to participate in this experiment.
>
> In the following experiment, you will learn which *components in the blood are indicators for a new illness and thus allow its diagnosis*. Please imagine that you are working in a medical laboratory. Recently there were several cases of this new illness. In order to *diagnose this illness*, you use a new procedure for investigating blood samples.
>
> You will see the sample of a test person on the screen at a time. Afterwards please indicate whether *the blood components in this blood sample indicate a disease*, thus whether the test person has gotten sick or not. Perhaps you will not be perfectly sure with your decision, therefore an 8-point answer scale is available, ranging from "1 = definitely not ill" to "8 = definitely ill". After you delivered your rating, you will receive feedback whether the disease occurred or not. At the beginning you will have to guess but after a while you will learn *what indicators for a disease are.* It doesn't matter how fast you answer, therefore, take the time you need. If you have any question, please ask now. Otherwise, please start the experiment by clicking on the "next" button.

The rating question asked "Are the components in this blood sample indicating a disease?".

Results and Discussion

Figure 2 displays the mean rating for the relevant stimuli in the test stage. The left panel shows the results of the group with a causal cover story. The group with a predictive cover story is shown on the right side. Again, within each panel, the black bars represent the mean rating of the previously trained stimulus (AB for the Added

Cue task on the left side and EFG for the Removed Cue task on the right side) whereas the white bars represent the rating of the corresponding test stimulus (ABX and EF/EG/FG, respectively).

In both groups the ratings of test stimuli were lower than those of the reinforced training stimuli. As in Experiment 1, the decrement seemed to be larger for the Removed Cue than for the Added Cue task. This was confirmed by a 2 x 2 x 3 x 2 ANOVA including the within-participants factors cue (trained vs. test), task (added vs. removed) and block (1-3) and the between-participants factors cover story (causal vs. predictive). There were main effects of cue (F (1, 41) = 61.00, $p < .001$, $\eta_p^2 = .598$), and a Cue*Task interaction (F (1, 41) = 4.09, $p < .05$, $\eta_p^2 = .091$). None of the other effects was significant with the main effect of instruction having the highest F of 2.973 ($p > .09$, $\eta_p^2 = .068$).

Fig. 2 about here

Again, we found asymmetrical generalization decrements as in Experiment 1. In contrast to Pineño et al. (2005), the cover story had no effect on the amount of overshadowing in the Removed Cue task and therefore on the asymmetry of the decrements. Therefore, the divergence between Experiment 1 and the previous experiments of Thorwart and Lachnit (2009) can not be traced back to the introduction of a causal cover story in the former.

General Discussion

In two experiments we observed asymmetrical generalization decrements. This is in contrast to our previous finding (Thorwart & Lachnit, 2009) but in line with other studies (e.g., Wheeler et al, 2006). This indicates that stimulus processing and representation differed between the experiments. As described above, asymmetrical decrements are in line with elemental models (Wagner, 2003; Harris, 2006) whereas the previously observed symmetry needs a configural explanation (Pearce, 1994).

However, neither stimulus duration, nor the kind of motion, nor the cover story was able to account for changes in processing. None of these manipulations was sufficient in the present experiments to result in symmetrical decrements, as observed in Thorwart and Lachnit. The divergence is in particular astonishing comparing the present Experiment 2 with Experiment 3 of Thorwart and Lachnit. The same kind of stimuli and training as in the present Experiment 2 was used in Group Fixed of Experiment 3. Despite this similarity, the contrasting results between the experiments suggest different kind of stimulus processing.

Searching for an explanation, the design of the test stage comes to the fore. In Thorwart and Lachnit (2009) we discussed the possibility that, in principle, each of the discriminations could be solved by relying only on one of the colored dots[1]. There is no need to use all the components presented in the training stimuli or to remember combination of the dots. This is unimportant for the Added Cue task, as all components of the trained stimulus are also part of the test stimulus and therefore participants in the test trial have all information they possibly used during initial discrimination learning. In contrast, if indeed a participant had relied on only one dot, this participant would not predict any disease in trials where this "relevant" component is removed in test, i.e. the decrement should be at maximum. On the other hand, removing “irrelevant" components would result in no decrement at all. Accounting for this possibility, in the present experiments the same new colored dot was added in the three test trials of the Added Cue task but each of the three trained components was removed in one test trial for the Removed Cue task, similar to Experiment 2 of Thorwart and Lachnit. In contrast, in Experiment 3 of Thorwart and Lachnit only two test blocks were presented and in each, the same, for each participant randomly selected dot was removed. The analyses up to now were based on the average over test trials and participants. The differing test for the Removed Cue task should then make no difference, as the relevance of the removed cue was balanced within subject in the present experiments and randomized between subjects in the previous Experiment 3. In Experiment 1 of Thorwart and Lachnit, not all binary combinations of the five-dots training stimuli were tested but only five (AB, BC, CD,

DE, AE) of them. But again, averaged across participants and trials, this should make no difference. Nevertheless, because of the surprising and divergent results in the present experiments, we analyzed the single trials of the test stage in order to see whether or not this more detailed inspection may provide clues concerning their origins.

Reanalyzes of the experiments: a more detailed inspection of the test stages

In order to investigate the impact of removing different dots, we have to identify the appropriate test trials for each participant. We determined the lowest and highest rating for each trial type, assuming that for example the trial with the lowest rating for the test stimulus of the Removed Cue task is the trial where the most relevant component was removed. The minimal and maximal ratings of the other trial types and all trial types of Experiment 3 were also identified to account for random fluctuation.

Figure 3 and 4 display the mean ratings for the minimal and maximal ratings in the present and the previous experiments, respectively. In Figure 3, it seems that the overall asymmetry in both present experiments is based on very low minimal ratings for removed test stimuli. As assumed, the trained stimuli and the test stimulus for the Added Cue task (left bars) varied only slightly and were rated on a high level throughout the test stage (i.e. even the minimal ratings are quite high). It is the rating of the Removed Cue test stimulus (right bar) that varied. For the minimal ratings, the rating of the Removed Cue test stimulus is clearly lower than the rating of the Added Cue test stimulus; for the maximal rating, it is at least descriptively higher in both experiments.

Fig. 3 about here

We reanalyzed the test stage of Experiment 1 using only the minimal or maximal rating in two 2 x 2 x 2 x 2 ANOVAs including the within-participants factors

cue (trained vs. test) and task (added vs. removed) and the between-participants factors stimulus duration (2 sec vs. 8 sec) and motion (grouped vs. separated). The Cue*Task interaction was significant for the minimal rating (F (1, 122) = 19.62, p < .001, η_p^2 = .139) but not for the maximal ratings (F <1). The cue x task x cover story ANOVAs for Experiment 2 revealed the same pattern: a significant Cue*Task interaction in the minimal rating (F (1, 41) = 7.93, p < .008, η_p^2 = .162) but not in the maximal rating (F < 1, for the Cue*Stage*Cover Story interaction F was 1.7).

Figure 4 displays the minimal and maximal ratings of Thorwart and Lachnit (2009)[2]. It seems that ratings in Experiment 1 were influenced by the relative relevance: for the minimal ratings, removing components produced a larger decrement than adding; for the maximal ratings, the opposite seemed to be the case. Experiment 2 revealed a minor asymmetry for the minimal ratings whereas the decrements were symmetrical in the maximal ratings. In Experiment 3, the differences between minimal and maximal ratings were lacking as expected based on the different test stage.

Fig. 4 about here

The cue x task x motion-ANOVAs for Experiment 1 of Thorwart and Lachnit (2009) revealed that there was a Cue*Task interaction for both minimal (F (1, 77) = 6.204, p < .016, η_p^2 = .075) and maximal ratings (F (1, 77) = 5.687, p < .021, η_p^2 = .069), confirming that the decrement for removing a cue was larger in the minimal ratings but smaller in the maximal ratings. In contrast, the Cue*Task interactions were not significant in Experiment 2 and Experiment 3 (all Fs < 1). There, the generalization decrements were symmetrical in both minimal and maximal ratings.

We now have different patterns of results. The minimal ratings were asymmetrical in Experiment 1 of Thorwart and Lachnit (2009) as well as in both current experiments, but not in Experiment 2 and 3 of Thorwart and Lachnit. For the maximal ratings, a switched asymmetry was observed in Experiment 1 of Thorwart

and Lachnit but symmetrical decrements in all other experiments. On the one hand, we found a similar pattern in Experiment 1 and the present experiments with the amount of asymmetry in favor of the Removed Cue task increasing from the maximal to the minimal rating. On the other side, we have Experiment 2 and 3 with no difference in symmetry between the minimal and maximal ratings. This further supports the above conclusions about causal and predictive cues. The differing cover stories between the experiment in Thorwart and Lachnit (2009) and the present experiments cannot account for the whole pattern of results. Instead, the influence of the removed dot differed between Experiment 1 and 2 in Thorwart and Lachnit.

Simulations with different saliences

Within all models discussed so far, the amount of responding to each component compared to other components being presented simultaneously during training depends on their relative saliences. In our (and as far as we can see, also others) previous simulations, saliences were identical for all components. By contrast, in the following we used differing saliences to simulate the "relevance" of one component[4]. Our aim was to look for reasonable values of saliences that allow the models under consideration to fit the empirical data.

The results of Experiment 2 of Thorwart and Lachnit (2009), i.e. symmetrical decrements in average as well as in minimal and maximal ratings, are what the Pearce model predicts without any further assumptions on saliences. The Harris Model can predict symmetry across trials, if one makes the additional assumption that the saliences of the stimuli are in general smaller for the Added Cue than for the Removed Cue task. But this overall symmetry is then due to switched asymmetries in the single minimal and maximal ratings – what is not the case in Experiment 2. The Replaced Elements Model can only predict symmetry assuming that nothing is learned about the removed dot, i.e. that its salience is zero. Beside the relatively unreasonable assumption that all participants learned nothing about one color and therefore have only small decrements in the maximal rating of the Removed Cue task, the Replaced Elements Model still predicts asymmetrical decrements averaged across trials.

Thus, only the Pearce Model seems to be valid so far. The question then is whether the Pearce Model can deal with the remaining results or whether these results are only in line with the elemental models, enforcing us to assume a switch in processing. In the following, simulations for three dots are described, but we confirmed the relevant predictions with five dots, too. We set the saliences of the three components of the Removed Cue task, E, F, and G, to .25, .5, and .75, respectively. The following analyses, however, do again not depend on these specific values (e.g., .7, .8, and .9 work, too). Simulations with each component removed were computed by permutation of the three values over stimuli (.25 .5 .75; .25 .75 .5; .5 .75 .5, with G removed). If we use these parameters for the Added Cue task and allow them to permute over training (A, B) as well as the added component (X) in the same way, the Pearce Model predicts the decrements presented in the upper panel of Figure 5. The white bars show the predicted decrements for the Removed Cue task with differently relevant components removed. The right bar (salience=.25) corresponds to the maximal rating, the bar with salience of the removed dot = .75 corresponds to the minimal rating and the most left bar to the rating averaged across all test trials. In the black bars, the average over the three permutations of the Added Cue task is computed because the minimal as well the maximal ratings are also averaged across all participants and therefore across added dots with high or low impact on generalization. The simulations illustrate that the Pearce Model predicts asymmetric generalization decrements for both trials where the most and the least salient dot is removed. However, the asymmetry switches: the decrement in the Removed Cue task is smaller than in the Added Cue task if the least salient component is removed, whereas the opposite is the case for the most salient component removed. The averaged decrements are symmetrical. Thus, the Pearce Model is able to account for the data observed in Experiment 1 of Thorwart and Lachnit where switched asymmetries were reported for the minimal and maximal ratings as well as symmetrical decrements averaged across trials.

Fig. 5 about here

However, the Pearce Model can also predict switched asymmetries in the minimal and maximal ratings that result in an *a*symmetry averaged across trials. In the previous simulations, the same permutations were used in the Added Cue and Removed Cue task. Then, however, in the Removed Cue task all "participants", i.e. permutations, had the same saliences for their training cues E, F, and G, whereas the saliences for the training cues in the Added Cue task, A and B, differed between participants. Some had .25 and .75, some .25 and .5, some .75 and .5. In order to adjust the simulations in this point, saliences of A and B were set to .25 and .75 (or .75 and .25) for all subjects, while X still had either a low (.25), or a middle-sized (.5), or a high impact (.75) on generalization in the test stage. The lower panel of Figure 5 shows the predictions based on these assumptions. Because of reduced decrements for the Added Cue task, the asymmetry is also reduced for the case where the least salient component is removed (salience = .25) whereas the asymmetry increases in the case with the most salient component removed (salience=.75). This leads to an asymmetry averaged over the trials. Thus, the overall asymmetry in the present experiments as well as the asymmetry in the minimal ratings can be explained by the Pearce model. The missing significant asymmetry in the maximal ratings may be a ceiling effect: as we used the maximal rating for all kind of stimuli, the decrement in the Added Cue task was already small which resulted in a conservative test for any asymmetry.

In Experiment 3 of Thorwart and Lachnit (2009), only two test blocks were presented and in each, the same component was removed, and this was randomly selected for each participant. Therefore, we cannot observe the influence of different relevant dots removed. But as shown above, the Pearce Model can handle symmetry across test trials both due to symmetry in the minimal and maximal ratings as in Experiment 2 and due to switched asymmetries as in Experiment 1.

Conclusion

The results of the present experiments as well as the reanalyzes of Experiment 1 of Thorwart and Lachnit (2009) are in line with predictions of the configural model of Pearce (1994) assuming different saliences for the components. The results of Experiment 2 of Thorwart and Lachnit can also be explained by the Pearce Model, if one assumes no differences in the saliences. Their influence in Experiment 3 cannot be investigated because of a different test stage, but the averaged results are also in line with Pearce's Model. In sum, elemental models are not necessary so far.

If these post-hoc simulations are truly explaining the results, some factors in the experiments must control the importance of the salience and the amount of different learning about each dot. Experiment 1 and Experiment 2 of Thorwart and Lachnit (2009) differed in the number of dots and the stimulus arrangement. Experiment 1 used two- and five-dots compounds, whereas Experiment 2 used two- and three-dots compounds. Hence, a first idea could be that differing saliences of the dots are not as relevant if there are fewer components to compete for attention within each trial. This explanation is of course in conflict with the present experiments that also used two- and three-dots compounds, but revealed results captured only by simulations where the saliences differed.

A similar explanation concerns the number of differently colored dots used throughout the experiment. In the present experiments we had to use 10 different colors throughout training because the generalization task was implemented as within-participants factor. In Experiment 1 and 2 on the other hand, the generalization task was a between-participants factor. Because of the higher number of dots per compound in Experiment 1, 10 colors had to be discriminated in the Group Removed Cue, while in Experiment 2 only 6 colors occurred throughout the training of Group Removed Cue. This reduced the need to selectively attend and to learn only about the most salient dots. Furthermore, the dots were not moving in Experiment 2 but presented fixed on positions randomly distributed over the screen. The initial reason for this was to investigate whether the motion per se may have had a configural impact. As both experiments in Thorwart and Lachnit (2009) resulted in symmetrical

decrements, we had to reject this claim. In light of the new analyses, motion might not have influenced the kind of stimulus processing but enhanced the effects of differing saliences by drawing on attention – even if the motion was irrelevant for solving the discrimination. In Experiment 2, with fewer and fixed dots, participants probably had enough capacity and time to attend and learn alike about all colors presented.

The decrements' asymmetry in favor of the Removed Cue task increased from the maximal to the minimal ratings in Experiment 1 of Thorwart and Lachnit (2009) as well as in the present experiments. However, averaged over all ratings, the first experiment resulted in symmetrical decrements whereas the latter experiment resulted in asymmetrical decrements. At the moment, we cannot provide an explanation for this quantitative difference. It also difficult to figure out the remaining divergence between Experiment 3 of Thorwart and Lachnit and Experiment 2 of the present experiments, especially as we do not know how the symmetrical decrements are generated in the former.

To sum up, when analyzed more precisely even apparently contradicting results of similar experiments can all be explained by one model. In the case of generalization decrements, we especially have to consider the relative saliences of each component present in compound. Then, the configural model of Pearce (1994) can account for our results. To be fair, the Harris Model is also quite good as it can simulate e.g. the switched asymmetries - but not the results of Experiment 2 of Thorwart and Lachnit (2009).

Whether or not this configural explanation can also account for the asymmetrical decrements of the studies from the others lab is difficult to say. None of them explicitly looked for differential effects of removing different components of the trained stimulus but counterbalanced the cue removed instead. Furthermore, more direct measurements of the salience/relevance of the single components for each participant are necessary in future investigations in order to provide external validation of our simulations.

References

Brainard, D. H. (1997). The Psychophysics Toolbox, *Spatial Vision,* 10, 433-436.

Brandon, S. E., Vogel, E. H., & Wagner, A. R. (2000). A componential view of configural cues in generalization and discrimination in Pavlovian conditioning. *Behavioural Brain Research*, 110, 67–72.

Glautier, S. (2004). Asymmetry of generalization decrement in causal learning. *Quarterly Journal of Experimental Psychology,* 57B, 315–329.

Glautier, S., Redhead, E., Thorwart, A., & Lachnit, H. (in press). Reduced Summation with Common Features in Causal Judgements. *Experimental Psychology.*

González, F., Quinn, J. J., & Fanselow, M. S. (2003). Differential effects of adding and removing components of a context on the generalization of conditional freezing. *Journal of Experimental Psychology: Animal Behavior Processes,* 29, 78-83.

Harris, J. A. (2006). Elemental representations of stimuli in associative learning. *Psychological Review,* 113, 584–605.

Harris, J. A., Livesey, E. J., Gharaei, S. & Westbrook, R. F. (2008). Negative patterning is easier than a biconditional discrimination. *Journal of Experimental Psychology: Animal Behavior Processes,* 34, 494-500.

Kinder, A., & Lachnit, H. (2003). Similarity and discrimination in human Pavlovian conditioning. *Psychophysiology,* 40, 226-234.

Lachnit, H., Schultheis, H., König, S., Üngör, M., & Melchers, K. G. (2008). Comparing elemental and configural associative theories in human causal learning: A case for attention. *Journal of Experimental Psychology: Animal Behavior Processes*, 34, 303-313.

Lachnit, H., Thorwart, A., Schultheis, H., Uengoer, M., Lotz, A., & König, S. (2009). *Indicators of early and late processing reveal the importance of within-trial-time for theories of associative learning*. Submitted manuscript.

Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review,* 82, 276-298.

Melchers, K. G., Shanks, D. R., & Lachnit, H. (2008). Stimulus coding in human associative learning: Flexible representations of parts and wholes. *Behavioral Processes,* 77, 413–427.

Myers, K. M., Vogel, E. H., Shin, J., & Wagner, A. R. (2001). A comparison of the Rescorla-Wagner and Pearce models in a negative patterning and a summation problem. *Animal Learning & Behavior,* 29, 36-45.

Palmer, S. E. (1999). *Vision science: Photons to phenomenology*. Cambridge, MA: MIT Press.

Pearce, J. M. (1994). Similarity and discrimination: A selective review and a connectionist model. *Psychological Review,* 101*,* 587-607.

Pearce, J. M., Aydin, A., & Redhead, E. S. (1997). Configural analysis of summation in autoshaping. *Journal of Experimental Psychology: Animal Behavior Processes,* 23, 84-94.

Pearce, J. M., & Bouton, M. E. (2001). Theories of associative learning in animals. *Annual Review of Psychology,* 52, 111-139.

Pearce, J. M., & Hall, G. (1980). A model for Pavlovian conditioning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review,* 87, 532-552.

Pineño, O., Denniston, J. C., Beckers, T., Matute, H., & Miller, R. R. (2005). Contrasting predictive and causal values of predictors and of causes. *Learning & Behavior*, 33, 184–196.

Redhead, E. S., & Pearce, J. M. (1995). Similarity and discrimination learning. *Quarterly Journal of Experimental Psychology*, 48B, 46-66.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current theory and research* (pp. 64-99). New York: Appleton-Century-Crofts.

Schultheis, H., Thorwart, A., & Lachnit, H. (2008a). HMS: A MATLAB simulator of the Harris model of associative learning. *Behavior Research Methods,* 40*,* 442-449.

Schultheis, H., Thorwart, A., & Lachnit, H. (2008b). Rapid-REM - A MATLAB simulator of the replaced-elements model. *Behavior Research Methods,* 40*,* 435-441.

Thorwart, A., Glautier, S., & Lachnit, H. (2009). *Addition of a common cue does not affect feature-negative discriminations in eyeblink conditioning and contingency learning in humans. S*ubmitted for publication.

Thorwart, A. & Lachnit, H. (2009). Symmetrical generalization decrements: configural stimulus processing in human contingency learning. *Learning & Behavior*, 37, 107-115.

Thorwart, A., Schultheis, H., König, S. & Lachnit, H. (2009). ALTSim: A MATLAB simulator for current associative learning theories. *Behavior Research Methods,* 41, 29-34.

Wagner, A. R. (2003). Context-sensitive elemental theory. *Quarterly Journal of Experimental Psychology,* 56B, 7–29.

Wheeler, D. S., Amundson, J. C., & Miller, R. R. (2006). Generalization decrement in human contingency learning. *Quarterly Journal of Experimental Psychology,* 59, 1212–1223.

Wasserman, E. A., & Miller, R. R. (1997). What's elementary about associative learning? *Annual Review of Psychology*, 48, 573-607.

Author Note

Anna Thorwart and Harald Lachnit, Department of Psychology, Philipps-Universität Marburg, Germany. This research is part of the doctoral thesis of Anna Thorwart. She was supported by the DFG (Deutsche Forschungsgemeinschaft) graduate program NeuroAct (DFG 885) and Harald Lachnit by grant La 564/19-1.

We would like to thank an anonymous reviewer for valuable and constructive comments on an earlier version.

Correspondence concerning this article should be addressed to Anna Thorwart, Department of Psychology, Philipps-Universität Marburg, Gutenbergstraße 18, 35032 Marburg, Germany (email: anna.thorwart@staff.uni-marburg.de).

Notes

[1] We would like to thank Mike Le Pelley for drawing our attention to this issue.

[2] The criterion for including participants was the same as in the present experiment.

[3]For the Pearce Model, the computation of the similarity was based on Eq. (9) in Pearce (1994, p. 600). The Replaced Elements Model as described by Wagner (2003, 2008) does not include any salience parameters for the components. In the following and after consultation with A. Wagner (personal communication, July 16, 2009), we assume that a highly salient CS will activate all of its elements (context-independent and dependent) more strongly than a less salient component.

Figure Captions

Figure 1. Mean ratings of the reinforced trained stimuli (AB in the Added Cue task and EFG in the Removed Cue task) and the corresponding test stimuli (ABX or EF/EG/FG, respectively) in the test stage of Experiment 1 for the four conditions.

Figure 2. Mean ratings of the reinforced trained stimuli (AB in the Added Cue task and EFG in the Removed Cue task) and the corresponding test stimuli (ABX or EF/EG/FG, respectively) in the test stage of Experiment 2 for the two instructions.

Figure 3. Minimal and maximal ratings of the reinforced trained stimuli (AB in the Added Cue task and EFG in the Removed Cue task) and the corresponding test stimuli (ABX and EF/EG/FG, respectively) in the test stage of Experiment 1 and 2, averaged over groups.

Figure 4. Minimal and maximal ratings of the reinforced trained stimuli and the corresponding test stimuli for the Added Cue and the Removed Cue task in the test stage of Experiment 1-3 of Thorwart and Lachnit (2009).

Figure 5. Simulation of the Pearce Model. Both panels display the predicted decrement between the reinforced trained stimuli and the corresponding test stimuli in the Added Cue (black bars) and the Removed Cue task (white bars). For further explanations see text.

fig1

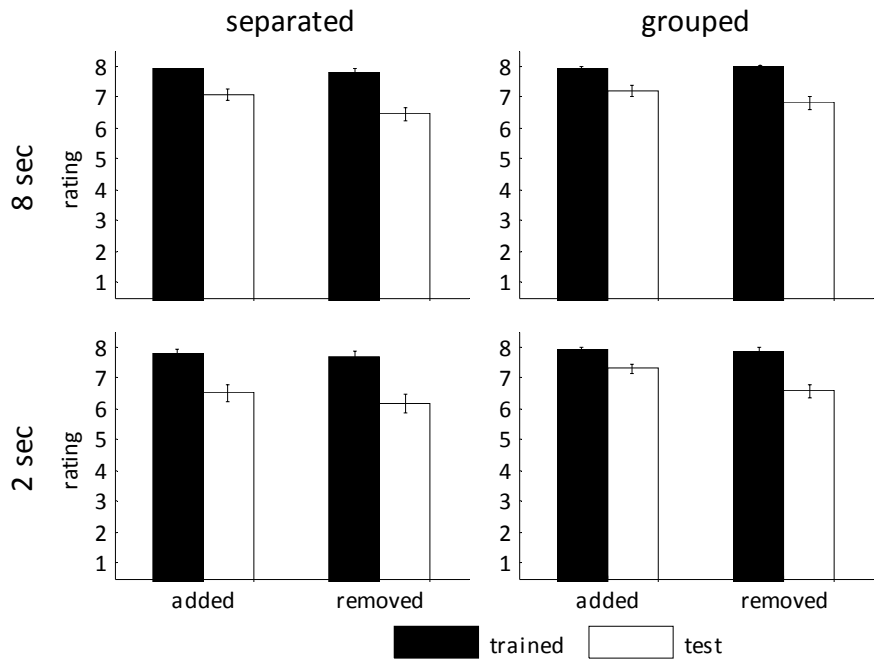

separated
grouped
8 sec
2 sec
rating
1
2
3
4
5
6
7
8
added
removed
trained
test

fig2

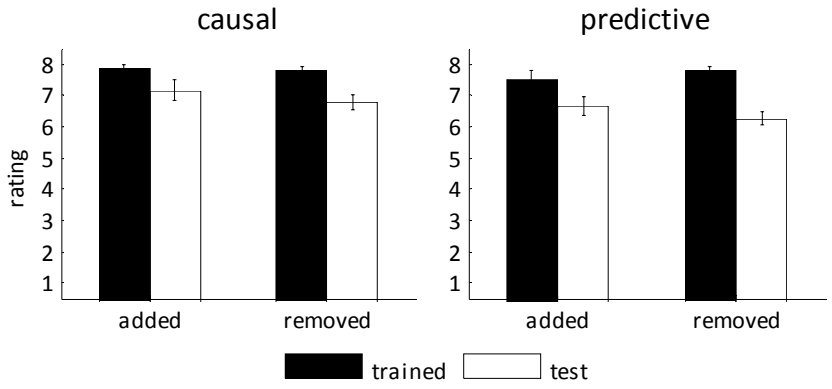

causal
predictive
rating
8
7
6
5
4
3
2
1
added
removed
added
removed
trained
test

fig3

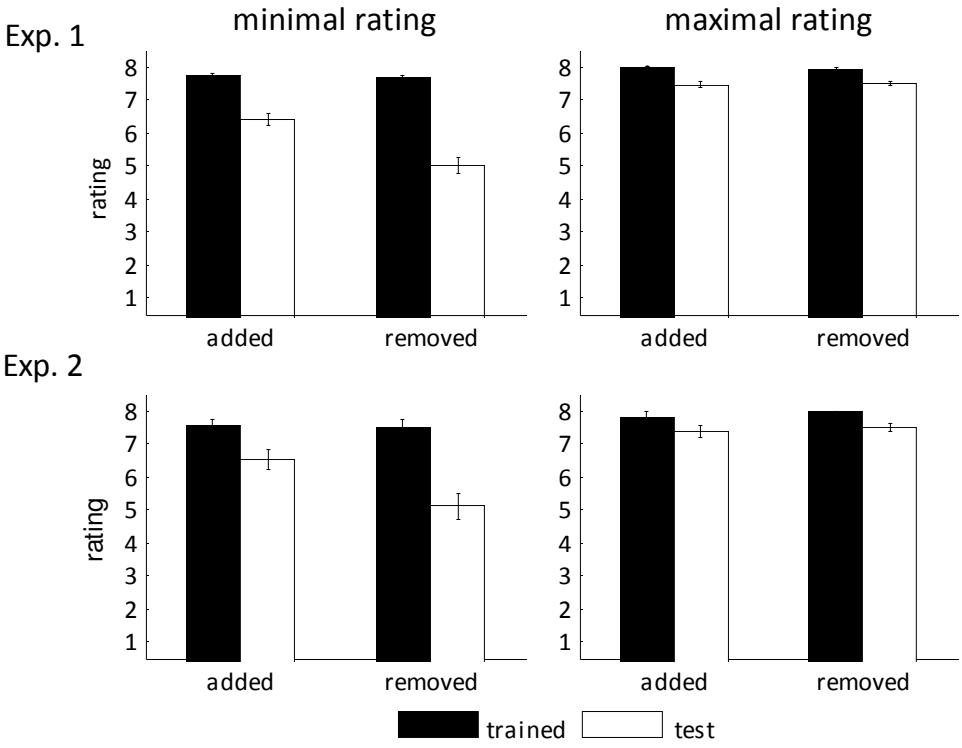

Exp. 1
minimal rating
maximal rating
rating
8
7
6
5
4
3
2
1
added
removed
Exp. 2
trained
test

fig4

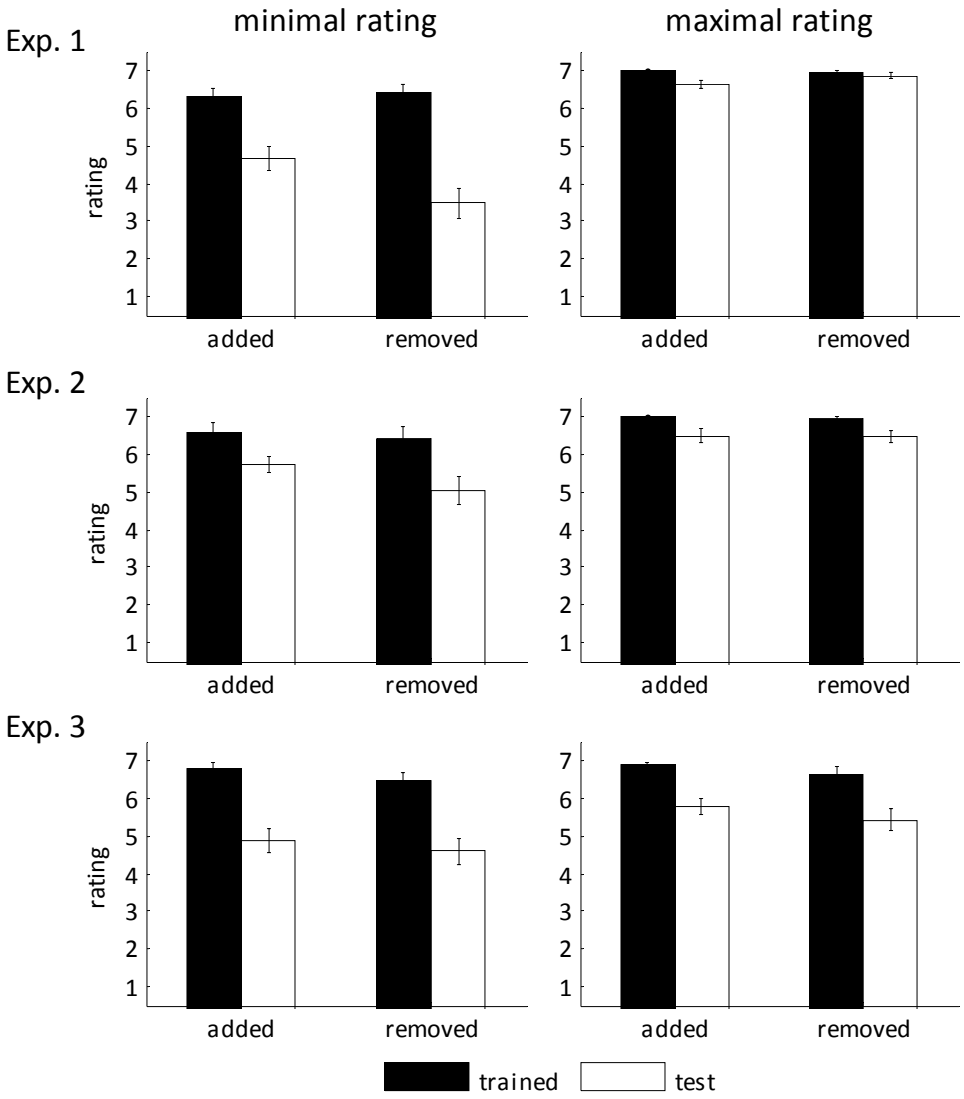
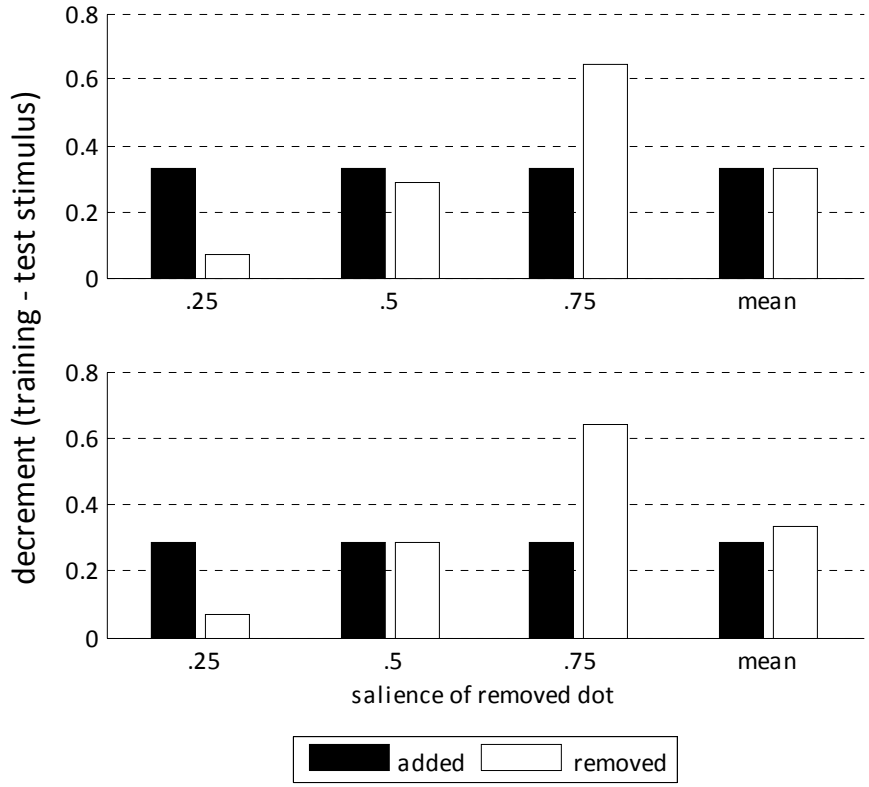

minimal rating
maximal rating
Exp. 1
Exp. 2
Exp. 3
rating
1
2
3
4
5
6
7
added
removed
trained
test

fig5


0.8
0.6
0.4
0.2
0
.25
.5
.75
mean
decrement (training - test stimulus)
salience of removed dot
added
removed

Curriculum Vita : Anna Barbara Thorwart

born 07 February 1981 in Nairobi, Kenya; Citizenship: German

office
Department of Psychology
Philipps-Universität Marburg
Gutenbergstraße 18
35032 Marburg, Germany

Phone: (+49) 06421 2823715
Fax: (+49) 06421 2826621
email: anna.thorwart@staff.uni-marburg.de
web page: http://www.staff.uni-marburg.de/~thorwara

home
Gisselbergerstraße 2
35037 Marburg, Germany

Phone: (+49) 06421 917021

Higher Education

2000 -2006	Diplom in Psychology, Philipps-Universität Marburg *Grade*: "sehr gut" (with distinction) *Thesis*: "Elemental and configural processing of complex stimuli during human classical conditioning" (Supervisor: Prof. Dr. Harald Lachnit)
2007 – (2010)	Ph.D. in graduate program "NeuroAct", Philipps-Universität Marburg *Supervisor*: Prof. Dr. Harald Lachnit (primary), Philipps-Universität, Marburg, and Prof. Dr. John Pearce, University of Cardiff *Thesis* : " Stimulus Representation and Processing in Human Associative Learning: An evaluation of current elemental and configural associative learning theories "

Professional Experience

June 2005 - July 2006	Research assistance, Department of Psychology, Associative Learning Group, Philipps-Universität Marburg
since August 2006	Research scientist and occasional lecturer, Department of Psychology, Associative Learning Group, Philipps-Universität Marburg

Related Experience

September– November 2003	Research Intern in the lab of Prof. Dr. Ruxandra Sireteanu, Max Planck Institute for Brain Research, Frankfurt
December 2003 -February 2004	Research Intern in the lab of Dr. Stefan Koelsch, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig
April - July 2007	Visiting PhD student, University of Southampton, Dr. Steven Glautier

Publications

Thorwart, A., Glautier, S. & Lachnit, H. (2009). *Addition of a common cue does not affect feature-negative discriminations in eyeblink conditioning and contingency learning in humans.* In preparation.

Lachnit, H., Thorwart, A., Schultheis, H., Uengoer, M., Lotz, A., & König, S. (2009). *Indicators of early and late processing reveal the importance of within-trial-time for theories of associative learning.* Submitted manuscript.

Thorwart, A. & Lachnit, H. (2009a). *Generalization decrements and salience: new data and more elaborated analyses*. *S*ubmitted manuscript.

Glautier, S., Redhead, E., Thorwart, A., & Lachnit, H. (in press). Reduced Summation with Common Features in Causal Judgements. *Experimental Psychology.*

Thorwart, A. & Lachnit, H. (2009b). Symmetrical generalization decrements: configural stimulus processing in human contingency learning. *Learning & Behavior*, 37, 107-115.

Thorwart, A., Schultheis, H., König, S. & Lachnit, H. (2009). ALTSim: A MATLAB simulator for current associative learning theories. *Behavior Research Methods*, 41, 29-34.

Schultheis, H., Thorwart, A., & Lachnit., H. (2008a). HMS: A MATLAB simulator of the Harris model of associative learning. *Behavior Research Methods*, 40, 442-449.

Schultheis, H., Thorwart, A., & Lachnit, H. (2008b). Rapid-REM: A MATLAB simulator of the replaced elements model. *Behavior Research Methods*, 40, 435-441.

Conference Contributions & Talks

Thorwart, A., Glautier, S., & Lachnit, H. (2009). Similarity does not affect Feature Negative discriminations in humans. Poster presented at the Associative Learning Symposium (XIII), Gregynog Hall.

Glautier, S., Redhead, E. , Thorwart, A., & Lachnit, H. (2008). Reduced summation with common features in causal judgements. Poster presented at EPA Meeting, Boston 2008.

Thorwart, A. & Lachnit, H. (2008). Symmetrie des Generalisationsdekrements und Reizverarbeitung beim Assoziativen Lernen. Talk presented at the "50. Tagung experimentell arbeitender Psychologen (TeaP)".

Thorwart, A., Lachnit, H., & Glautier, S. (2008). Stimulus Processing in Contingency Learning: Symmetry of Generalisation Decrements. Talk presented at the Associative Learning Symposium (XII), Gregynog Hall.

Thorwart, A. & Lachnit, H. (2007). Current associative theories predict CS processing but fail to explain outcome anticipation. Poster presented at Associative Learning Symposium (XI), Gregynog Hall.